AD_____

Award Number: W81XWH-11-1-0437

TITLE: A Forward Genetic Screening for Prostate Cancer Progression Genes

PRINCIPAL INVESTIGATOR: Tian Xu, Ph.D.

CONTRACTING ORGANIZATION: Yale University School of Medicine

New Haven, CT 06520-8047

REPORT DATE: October 2012

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| 01-10-2012 | Annual | 30 Sep 11-29 Sep 12 |

**4. TITLE AND SUBTITLE**
A Forward Genetic Screening for Prostate Cancer Progression Genes

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
DOD W81XWH-11-1-0437

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Tian Xu, Ph.D.
Betty Diamond

E-Mail: tian.xu@yale.edu

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Yale University School of Medicine

New Haven, CT 06520-8047

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

## 14. ABSTRACT

It has been recently established that up to a hundred genetic alterations occur during the development of prostate cancer. Some of these mutations likely cause cancer progression while others could be passengers with no causative effect in tumor development. We are implementing a forward genetics approach for prostate cancer research, by first inducing tumors by random genetic alterations and then mapping the recurring mutations in the prostate cancer. This approach efficiently identifies mutations that drive cancer development and progression. Currently, we are screening and further optimizing our mutagenesis approach for maximum efficiency. In addition, we are collecting tumors and mapping causative genes. Once identified these genes will be validated in human patient samples. Identifying which mutations are responsible for tumor progression and metastasis will likely identify new avenues for treating advanced stage cancer.

**15. SUBJECT TERMS**
Prostate Cancer, Progression, Metastasis, Forward Genetics, Transposon

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | 19b. TELEPHONE NUMBER *(include area code)* |
| U | U | U | UU | 50 | USAMRMC |

## Table of Contents

## INTRODUCTION:

Advanced prostate cancer is a common killer in the United States. The American Cancer Society estimates that more than 28,000 men will die of advanced stage prostate cancer in this year alone (1). Approximately 90% of all cancer deaths are due to metastasis, which is the spread of the cancer to distant sites. Although many cancer mutations have been identified, little is known about what drives cancer to progress and metastasize. Therefore, the creation of novel preclinical models to discover which mutations cause progression and metastasis would be of great value. Here, we describe a novel methodology, based on the piggyBac (*PB*) transposon, to induce and identify mutations that promote the progression and metastasis of solid tumors, and apply it to the analysis of prostate cancer. We have refined the highly efficient *PB* transposon system to induce mutations and added a reporter that allows us to visualize tumor development and the spread of cancer cells in live mice. By introducing this mutagenesis system into mice that develop early stage prostate tumors, we will be able to identify mutations that cause prostate tumors to advance and metastasize. Ultimately, we will then validate these identified mutations in human tumor samples. Identifying which mutations are responsible for tumor progression and metastasis will pave the way for understanding and treating advanced stage cancer.

## BODY:

**Hypothesis:** 1. The *PB*-based somatic mutagenesis will reveal genetic alterations responsible for progression of prostate cancer. 2. The genetic changes identified in mice are relevant for carcinogenesis of human prostate cancer.

**Objectives:**

1. **To establish the first forward genetic model for somatic mutagenesis of prostate tissue.**

   We have developed a somatic mutagenesis system for cancer gene discovery in the mouse based on the *PB* transposon. The high efficiency of *PB* transposition permits screens to be performed using few mutagenic transposons, offering the additional advantage of circumventing high background mutation rates in identifying causative mutations. As a result we have developed a highly efficient, low-copy *PB* mutagenesis system (PB-SMART) utilizing a seven-copy *PB* mutator called *Luc-PB[mut]7*, which can also monitor mutagenesis by luciferase expression. This system has been described in our recent publication where we demonstrate the utility of the system in whole animal and skin targeted mutagenesis (2, appendix pg. 39). We have also used this system to conduct a low-copy transposon mutagenesis screen for melanoma (appendix, pg.11). Exome sequencing has revealed that somatic mutations in human melanomas greatly exceed those in other cancers, making it a challenge for predicting causative genes from the large collection of mutations. In the attached paper, we demonstrate that low-copy transposon mutagenesis in mice offers a powerful and complementary way to decipher which alterations actually drive human melanoma. Currently, we are utilizing PB-SMART to screen for genes that drive prostate cancer progression as detailed below.

2. **To identify genes whose activation or inactivation may cause progression of prostate cancer.**

   We began screening in the *Nkx3.1*$^{+/-}$ background with the PB-SMART system and are currently aging the mice and observing them for tumor development. Specifically we have crossed *Nkx3.1*$^{+/-}$ mice into the *Luc-PB[mut]7/ ACT-Stop-PBase/ PB4-Cre+* background. Currently we have eight mice in this *Nkx3.1*$^{+/-}$ screening group that have surpassed a year of age and none

have developed signs of prostate tumor development. Ideally we would like to increase the incidence of tumor formation before a year to enhance our screening efficiency.

We have found that body-wide mutagenesis can be highly effective in inducing cooperating mutations with an organ-specific sensitizing mutation as demonstrated in the melanoma screen (appendix, pg. 11). As a result, we compared the efficiency of transposon mobilization in the anterior prostate with transposase expression restricted by the PB4 promoter (*Luc-PB[mut]7/ACT-Stop-PBase/ PB4-Cre+)* versus the *ACT* promoter (*Luc-PB[mut]7/ACT-PBase*). We have found that mutagenesis with *ACT-PBase* where transposon mobilization occurs in early development rather than in response to androgen from the PB4-CRE results in higher transposition efficiency (9-30 fold) in the anterior prostate (Fig. 1). Consequently, we are breeding cohorts with the *Luc-PB[mut]7/ACT-PBase* genotype to increase mutagenesis efficiency in the prostate.

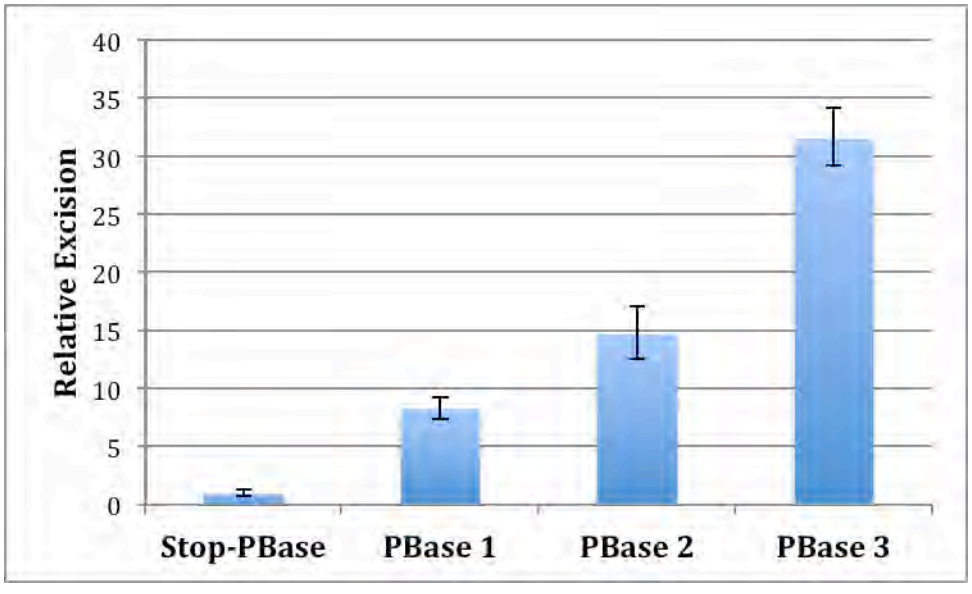


**Figure 1. *ACT-PBase* drives more efficient transposon mobilization than *Act-Stop-PBase/PB4-CRE.*** Quantitative PCR is used to measure relative excision of the *PB* transposon in DNA extracted from the anterior prostate. Stop-PBase represents the *Luc-PB[Mut]7/Act-Stop-PBase/PB4-CRE* genotype and PBase represents the *Luc-PB[Mut]7/Act-PBase* genotype.

We have also decided to increase the efficiency of the screen by using a more sensitized genetic background. The tumor suppressor *PTEN* is inactivated in approximately 40% of human prostate cancer (3, 4). Mice with prostate-specific deletion of the tumor suppressor *Pten* develop early benign cancer, prostatic intraepithelial neoplasia (PIN) that can progress to high-grade adenocarcinoma with long latency, but without progression to frank metastasis (5-8). Importantly, this phenotype can be pushed to metastatic disease with as little as one other genetic mutation (9). We have introduced the PB-SMART system into the prostate specific $Pten^{-/-}$ background to increase the chances of identifying mutations that drive progression to metastatic disease. Thus far we have observed 3/8 (37.5%) control mice (*$Pten^{lox/lox}$/ ACT-Stop-PBase/ PB4-Cre or $Pten^{lox/lox}$/ Luc-PB[mut]7/ PB4-Cre)* and 8/13 (61.5%) experimental mice (*$Pten^{lox/lox}$/ Luc-PB[mut]7 / ACT-Stop-PBase/ PB4-Cre)* that have developed prostate adenocarcinoma in a cohort of mice that are older than one year. We plan to continue to age additional mice from these genotypes to see if the penetrance is increased and latency is decreased with the addition of PB-SMART. We have observed aggressive prostate carcinoma with local invasion, but have failed to observe dissemination into the lumbar lymph nodes or other tissues (Fig. 2). We are currently mapping transposon insertions from the *$Pten^{lox/lox}$/ Luc-PB[mut]7 / ACT-Stop-PBase/ PB4-Cre* prostate tumors to identify recurrently affected genes or pathways. We believe that by increasing mutagenesis efficiency in the prostate with *ACT-PBase* we can further increase the penetrance, decrease the tumor latency, and induce macro-metastatic dissemination.

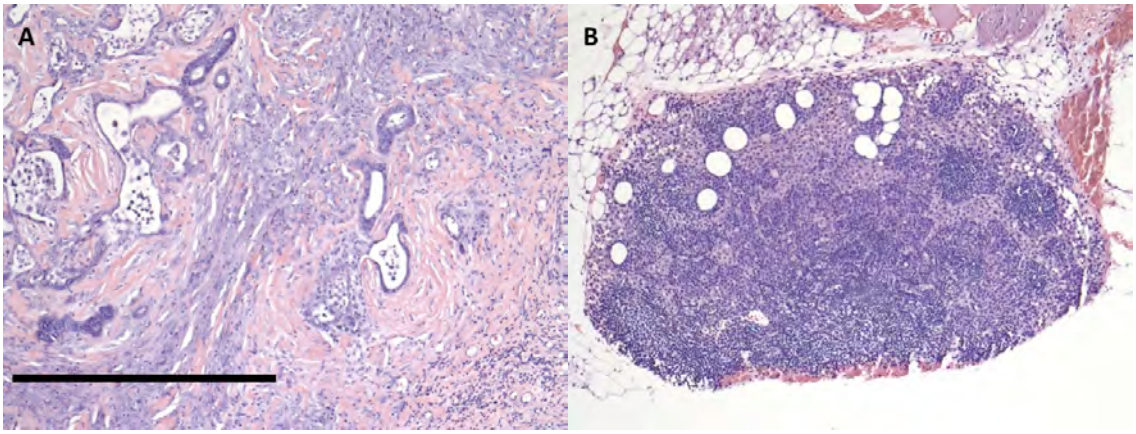


**Figure 2. Prostate adenocarcinoma develops in *Pten$^{lox/lox}$/ Luc-PB[mut]7 / ACT-Stop-PBase/ PB4-Cre* mice. A.** Representative haematoxylin and eosin (H&E) stained section of a prostate tumor from a *Pten$^{lox/lox}$/ Luc-PB[mut]7 / ACT-Stop-PBase/ PB4-Cre* . The sheets of invasive cells are indicative of adenocarcinoma. **B.** H&E stained section of a lumbar lymph node from the same mouse as in (A). Infiltration of prostate cancer cells is not evident. Scale bar, 500uM.

**3. To analyze the human prostate cancer tissues for possible abnormal expression and/or mutations of the candidate genes identified in the mouse model.**

We are currently mapping transposon insertions from our collection of prostate tumors to identify recurrently affected genes or pathways. We are also increasing the screening cohort to induce a larger collection of prostate tumors. Once we identify recurrently mutated genes or pathways we will examine them in human cancer prostate tissues with our collaborator Dr. Liu.

## KEY RESEARCH ACCOMPLISHMENTS:

- **Established the PB-SMART system for cancer gene discovery in the mouse.**
- **Implemented PB-SMART in the prostate to identify mutations that induce progression and metastasis.**
- **Identified and implemented strategies to increase the efficiency of the screen to reduce time and cost.**

## REPORTABLE OUTCOMES:

**Manuscripts:**
Establishing the PB-SMART system and implementing it for cancer gene discovery has resulted in one published article and one article under review.

Landrette SF, Cornett JC, Ni TK, Bosenberg MW, Xu T. *piggyBac* transposon somatic mutagenesis with an activated reporter and tracker (PB-SMART) for genetic screens in mice. PLoS One. 2011;6(10):e26650. Epub 2011/11/01.

Ni TK, Landrette SF, Bosenberg MW, Bjornson RD, Xu T. Low-copy *piggyBac* Transposon Mutagenesis in Mice Differentiates Genes Mutated in Human Melanomas. Science. Under Review.

**Cell lines:**
Thus far 2 cell lines have been established from the *Pten$^{lox/lox}$/ Luc-PB[mut]7 / ACT-Stop-PBase/ PB4-Cre* tumors.

**Animal Models:**
We have established the PB-SMART system (*Pten$^{lox/lox}$/ Luc-PB[mut]7 / ACT-Stop-PBase/ PB4-Cre*) for cancer gene discovery in the prostate.

**CONCLUSION:**

We have established the PB-SMART system and are implementing it for cancer gene discovery in the prostate. We isolated a collection of tumors, which we are using to identify causative mutations driving progression. Once identified, these driver genes and pathways will be analyzed in human prostate tumors to confirm their relevance to the human disease. Furthermore, we have identified strategies to increase the mutagenesis efficiency of PB-SMART in the prostate. Optimizing screening efficiency will increase the penetrance and progression of the prostate tumors, and as a result reduce time and cost. Thus, we are poised to identify mutations that drive prostate tumor progression and metastasis, which should contribute to a better understanding and treatment of advanced stage cancer.

**REFERENCES:**


1. American Cancer Society. Cancer Facts and Figures. Atlanta: American Cancer Society, 2012.
2. Landrette SF, Cornett JC, Ni TK, Bosenberg MW, Xu T. piggyBac transposon somatic mutagenesis with an activated reporter and tracker (PB-SMART) for genetic screens in mice. PLoS One. 2011;6(10):e26650. Epub 2011/11/01.
3. Pourmand G, Ziaee AA, Abedi AR, Mehrsai A, Alavi HA, Ahmadi A, et al. Role of PTEN gene in progression of prostate cancer. Urology journal. 2007;4(2):95-100. Epub 2007/08/19.
4. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. Cancer cell. 2010;18(1):11-22. Epub 2010/06/29.
5. Wang S, Gao J, Lei Q, Rozengurt N, Pritchard C, Jiao J, et al. Prostate-specific deletion of the murine Pten tumor suppressor gene leads to metastatic prostate cancer. Cancer cell. 2003;4(3):209-21. Epub 2003/10/03.
6. Chen Z, Trotman LC, Shaffer D, Lin HK, Dotan ZA, Niki M, et al. Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. Nature. 2005;436(7051):725-30. Epub 2005/08/05.
7. Trotman LC, Niki M, Dotan ZA, Koutcher JA, Di Cristofano A, Xiao A, et al. Pten dose dictates cancer progression in the prostate. PLoS biology. 2003;1(3):E59. Epub 2003/12/24.
8. Ma X, Ziel-van der Made AC, Autar B, van der Korput HA, Vermeij M, van Duijn P, et al. Targeted biallelic inactivation of Pten in the mouse prostate leads to prostate cancer accompanied by increased epithelial cell proliferation but not by reduced apoptosis. Cancer research. 2005;65(13):5730-9. Epub 2005/07/05.
9. Ding Z, Wu CJ, Chu GC, Xiao Y, Ho D, Zhang J, et al. SMAD4-dependent barrier constrains prostate cancer growth and metastatic progression. Nature. 2011;470(7333):269-73. Epub 2011/02/04.

## APPENDICES:

# Title: Low-copy *piggyBac* Transposon Mutagenesis in Mice Differentiates Genes Mutated in Human Melanomas

**Authors:** Thomas K. Ni[1], Sean F. Landrette[1], Robert D. Bjornson[2], Marcus W. Bosenberg[3], Tian Xu[1,4]*

**Affiliations:**

[1]Department of Genetics, Howard Hughes Medical Institute, Yale University School of Medicine, Boyer Center for Molecular Medicine, New Haven, Connecticut, United States of America

[2]Department of Computer Science, W.M. Keck Biotechnology Resource Laboratory, Yale Center for Genome Analysis, Yale University, New Haven, Connecticut, United States of America

[3]Department of Dermatology, Yale University School of Medicine, New Haven, Connecticut, United States of America

[4]Institute of Developmental Biology and Molecular Medicine, Fudan-Yale Center for Biomedical Research, School of Life Science, Fudan University, Shanghai, China

*To whom correspondence should be addressed. Email: tian.xu@yale.edu

**Abstract:** Exome sequencing has revealed that somatic mutations in human melanomas greatly exceed those in other cancers, making it a challenge for predicting causative genes from the large collection of mutations. Here we used low-copy *piggyBac* (*PB*) transposon mutagenesis in mice with conditionally activated *Braf* in melanocytes to screen for genes promoting melanoma development. Analysis of eleven *PB*-induced melanomas revealed on average four mutated genes per tumor, circumventing the problem of high background mutations. The two *bona fide* melanoma drivers relevant to patients with oncogenic *BRAF* mutations, *CDKN2A* and *MITF*, were identified, demonstrating the validity and efficiency of the screen. Notably, analysis of the identified genes in conjunction with human melanoma mutational data revealed two recurrently mutated networks involving *MAGI2* and *MAP3K1/2*, and disruption of *MAGI2* can transform human melanocytes. Overall, 34 out of 38 identified genes are mutated in human melanomas, providing invaluable instructive information for sorting through the large collection of mutations and differentiating genes important for melanomas. In summary, low-copy transposon mutagenesis in mice offers a powerful and complementary way to decipher relevant alterations for human conditions, especially those accompanied by high degrees of genetic variation.

**One Sentence Summary:** Melanomas induced with low mutation rates by *piggyBac* transposon mutagenesis help pinpoint genes important for human melanomas.

**Main Text:** Cancer is driven by the somatic acquisition of genetic alterations, and identification of mutated genes is critical for understanding the underlying mechanisms that drive cancer and designing therapeutics. With the improvements in sequencing power, sequencing of protein-coding regions has proven to be highly effective in identifying mutations in human cancers. This approach is particularly fruitful in identifying causative mutations for many types of cancers, which harbor on average 60-70 somatic mutations (*1-3*), and those mutations driving tumorigenesis can be distinguished by their higher-than-background mutation rates. However, it has remained difficult to predict causative mutations for cancers with elevated background mutation rates such as melanomas, which have ten-fold more mutations (*4-9*). Thus, additional approaches are needed to differentiate those genes important for melanomas from passenger mutations.

Since *piggyBac* (*PB*) and *Sleeping Beauty* (*SB*) cut-and-paste transposons were demonstrated to mobilize in mice (*10, 11*), transposon somatic mutagenesis screens have represented promising avenues for cancer gene discovery in conjunction with human cancer sequencing data (*12-14*). For example, *SB* insertional mutagenesis screens have been successfully conducted for colon and pancreatic cancer in mice (*15-17*). Moreover, the high efficiency of *PB* transposition permits screens to be performed using few mutagenic transposons, offering the additional advantage of circumventing high background mutation rates in identifying causative mutations. We have previously developed a highly efficient, low-copy *PB* mutagenesis system (PB-SMART) utilizing a seven-copy *PB* mutator called *Luc-PB[mut]7*, which can also monitor mutagenesis by luciferase expression (*14*). We used this system to

conduct a low-copy transposon mutagenesis screen for melanomas in mice and analyzed the mouse mutational data together with human melanoma sequencing data.

**Low-copy *PB* Transposon-Based Genetic Screen for Melanomas.** Human nevi and melanomas frequently harbor oncogenic $BRAF^{V600E}$ mutations (*18, 19*) as BRAF activation is an important initiating event, albeit insufficient for the genesis of melanoma. To sensitize mice to melanomas, we utilized a conditional knock-in oncogenic *Braf* mouse model, $Braf^{CA}$ (*20, 21*). $Braf^{CA}$ mice were bred with mice carrying the melanocyte-specific *Tyr-CreER* transgene (*22*) to enable tamoxifen-dependent activation of oncogenic $Braf^{V600E}$.

We performed a forward genetic screen for melanomas by introducing our low-copy *PB* somatic mutagenesis system into this double transgenic background (Fig. 1A). A mutagenesis cohort of mice which mobilize *PB* mutagenic transposons in the soma ($Braf^{CA}$*;Tyr-CreER;Luc-PB[mut]7;Act-PBase*) and a non-mutagenesis control cohort ($Braf^{CA}$*;Tyr-CreER;Luc-PB[mut]7* or $Braf^{CA}$*;Tyr-CreER;Act-PBase* littermates) were monitored for the development of melanomas. As early as four months after tamoxifen treatment of the mutagenesis cohort, we detected luciferase-expressing foci from otherwise unremarkable mouse coats (Fig. 1B). Subsequent removal of hair revealed skin tumors (Fig. 1C) that were confirmed to be malignant melanomas by histology (Fig. 1D and E). Overall, 8/17 mice in the mutagenesis cohort succumbed to melanoma within fifteen months, while the control cohort remained mostly melanoma-free (Fig. 1F), indicating that low-copy *PB* somatic mutagenesis can efficiently induce melanomas.

**The Nature of Somatic Insertional Mutations.** To verify that low-copy *PB* mutagenic transposons can interrogate proto-oncogenes and tumor suppressor genes, we undertook molecular characterization of several dozen insertion sites in melanomas. 60% of insertions

were located in transcription units or genes (hereafter referred to as insertional mutations) and the rest were found in intergenic regions. We focused on analyzing insertions that disrupted genes and discovered five major types of insertional mutations. One type of insertional mutation involved coding direction-oriented insertions at 5' regions of genes. We found that these insertions can upregulate gene expression via the Actin promoter in the transposon as seen in the case of *Mitf* (Figs. 1A and 2A). A second type of insertional mutation involved insertions directly into exons and was found to disrupt gene expression, as in the case of *Cdkn2a* (Fig. 2B). Many insertions were mapped to introns, and these represented three additional types of insertional mutations. In either orientation (coding or opposite), intron insertions could disrupt gene expression (Fig. 2C). Insertions into introns could also induce the expression of Actin promoter-driven chimeric transcripts when oriented in the coding direction as in the case of *Map3k2*, leading to elevated levels of truncated *Map3k2* products (Figs. 1A and 2D) (*14*). Our results indicate that the mutagenic *PB* transposon utilized in this screen can disrupt the functions of both proto-oncogenes and tumor suppressor genes.

**Mapping Insertional Mutations by Illumina Sequencing.** Illumina sequencing of PCR-amplified insertions has been reported to uncover greater genetic complexity within transposon-induced tumors (*23*). We performed Illumina sequencing to comprehensively map and analyze the insertional landscape in the eight low-copy *PB*-induced melanomas isolated from our mutagenesis cohort described above (*24*). In addition, we also analyzed three low-copy *PB*-induced melanomas isolated from other mutagenesis screens lacking $Braf^{V600E}$. This analysis revealed that few insertional mutations comprised of the vast majority of sequencing reads in each tumor. The preponderance of reads from these insertions reflects their clonal

abundance in the melanomas and thus makes them potential drivers. A total of 45 insertional mutations in 38 genes were identified in the eleven melanomas analyzed, corresponding to an average of 4.1 insertional mutations per tumor (table S1) (*24*). We randomly chose 15/45 and performed PCR across insertion sites to verify the abundance of the transposon insertions within tumors. In all cases, the prevalence of inserted alleles were validated (fig. S1).

**Recurrently Mutated Genes Identify Mutually Exclusive *Map3k1/2* and *Braf$^{V600E}$* Mutations in Melanomas.** Even though the number of insertional mutations per tumor is extremely low, we nevertheless identified *Mitf*, *Arhgef3*, *Map3k1*, *Map3k2* and *Rapgef2* as recurrently mutated genes (Table 1). This strongly argues that the identified mutations and genes are important for promoting melanomas. Indeed, *Mitf* is a known melanoma oncogene in humans (*25, 26*). Meanwhile, a convergence of three independent insertions in the coding direction in *Arhgef3*, which encodes a guanine nucleotide exchange factor for Rho family members, supports a recent finding from melanoma exome sequencing that Rho family GTPases may be novel melanoma oncogenes (*9*).

With a collection of five mutations altogether, *Map3k1* and *Map3k2*, components of the *Mapk* pathway, were the most frequently mutated genes found among *PB*-induced melanomas. Transposon insertions in *Map3k1* were clustered in introns 9 and 10 (Fig. 3A). The clustering of insertions in this region is not due to an enrichment of *PB* target sites (TTAA) and thus suggests a functional significance. Indeed, all these insertions are oriented in the coding direction and cause overexpression of the truncated *Map3k1* C-terminal product containing the kinase domain (Fig. 3B-E). Similar to *Map3k1*, insertions in *Map3k2* also overexpress the C-terminal kinase domain as described above (Fig. 2D).

Approximately 65% of human melanomas harbor oncogenic *BRAF* lesions. It is interesting whether the rest of patients without *BRAF* mutation also rely on MAPK pathway activation. It is known that MAP3K1/2 members can activate MAPK signaling independent of BRAF (*27, 28*), raising the possibility that MAP3K1/2 could substitute for BRAF in melanomas. The conditional *Braf* allele in our screen made it possible to examine whether *Braf* was activated in *Map3k1/2*-mutated melanomas. We found that $Braf^{V600E}$ and *Map3k1/2* mutations were mutually exclusive in all but one *PB*-induced melanomas (Fig. 3F and Table 1; segregation of these mutations could not be distinguished in one tumor because it contained both *Braf* activated and non-activated cells). These data support the notion that *Map3k1/2* can functionally substitute for oncogenic *Braf* to activate the Mapk pathway in promoting melanomas (Fig. 3G).

**Analysis of Mouse Mutational Data in Conjunction with Human Melanoma Sequencing Data.** High mutation rates in human melanomas make it difficult to distinguish causal alterations from passengers. Since a small collection of mutations has been identified in our low-copy *PB* mutagenesis screen for driving melanomas in mice, comparing our results with human melanoma mutational data could not only validate our screen but also provide important instructive information for differentiating melanoma-promoting genes in humans. Indeed, *CDKN2A* and *MITF*, the two highly prevalent, *bona fide* melanoma drivers relevant to patients with oncogenic *BRAF* mutations, were identified in our collection. Although only several *bona fide* melanoma drivers are known, other genes important for melanomas must exist in human exome sequencing data sets (*5*-*9*). The intersection between the small number genes identified in our melanoma-promoting mutagenesis screen in mice with the mutations

found in human exome sequencing data could be an informative approach to differentiate genes that are potentially important for human melanoma development. For example, one of the genes at this intersection was *MAGI2*. In addition, we interestingly found that *MAGI2* is involved in a protein interaction network in which many members are mutated in both our screen and human melanomas or in one of the data sets (Fig. 4A) (*5-9, 24*). We found that depletion of *MAGI2* (Fig. 4B-C) conferred anchorage-independent growth to human melanocytes stably expressing $BRAF^{V600E}$ in soft agar assays (Fig. 4D-E) (*25*). Collectively, these results provide proof-of-principle examples indicating that low-copy transposon mutagenesis in mice and intersectional analysis with human melanoma exome sequencing data are highly useful for distinguishing genes important for melanomas.

We compared genes from our screen to mutational data in human melanomas as reported by five exome sequencing studies as well as the Catalogue of Somatic Mutations in Cancer (COSMIC) and found extensive intersections (Fig. 5A) (*5-9, 29*). Including recurrently mutated genes, *CDKN2A*, *MITF* and *MAGI2*, 34 out of the 38 genes identified in our screen are mutated in human melanomas, and 28 of them are mutated in three or more samples (Fig. 5A-B and table S2). The strikingly high degree of intersection is, on one hand, not surprising because the low-copy transposon mutagenesis used in the mouse screen dictates that the mutations are relevant to the phenotype. On the other hand, this raises an interesting question regarding why these intersecting genes do not stand out in the analysis of human melanoma mutational data alone.

One possibility is that the high mutation rate in human melanomas obscures the causal mutations (Fig. 5C). We thus calculated the expected random mutation frequencies taking into

account the size of each gene and the average melanoma mutation rate of 650 mutations per tumor (*5-9*) and compared those with the detected mutation frequencies revealed by melanoma exome sequencing (Fig. 5D; expected *x-axis*, observed *y-axis*; orange line represents equal value of observed versus expected). Interestingly, only *CDKN2A*, a very small gene, is significantly mutated above the expected random mutation frequency (Fig. 5D). Moreover, even *MITF* is mutated much less frequently than expected by chance despite having been independently verified as a *bona fide* melanoma driver (Fig. 5D) (*25*). However, if the mutation rate were ten times lower, as found in other cancer types (*1-3*), most of these genes including *MITF* would have detected mutation frequencies higher than expected (Fig. 5D; green line). This illustrates why exome sequencing is a highly effective approach for identifying causal mutations in many cancers (*1-3*) and underscores how high mutation rates could complicate this approach. Indeed, after more than 150 melanomas sequenced, few predicted candidate drivers have been corroborated by independent sequencing studies (*5-9*). The low-copy transposon mutagenesis in mice in conjunction with the human exome sequencing data is thus particularly useful to differentiate human somatic mutations and identified a group of genes important for melanomas. This approach represents a powerful way for identifying genes relevant to diseases, especially those that are affected by high mutation rates.

**Figure Legends**

**Fig. 1.** Low-copy *PB* somatic mutagenesis-induced melanomas. (**A**) A marker-assisted breeding strategy generates fluorescent-eyed, grey-furred quadruple transgenic mice for low copy transposon screening with conditional *Braf* activation in melanocytes. (**B**) Focal luciferase expression (arrowhead) appears during the monitoring of *PB* mutagenesis. (**C**) Subsequent removal of the coat reveals a skin tumor beneath. (**D** and **E**) Histological analysis of skin tumors finds malignant melanomas (scale bar 50 µm), seen here invading through muscle (D) and staining positively for the S100 melanoma marker (E). (**F**) High melanoma incidence was observed in the mutagenesis cohort (red, *n* = 17, *Braf$^{CA}$;Tyr-CreER;Luc-PB[mut]7;Act-PBase*) compared to control cohort (blue, *n* = 17, *Braf$^{CA}$;Tyr-CreER;Luc-PB[mut]7* or *Braf$^{CA}$;Tyr-CreER;Act-PBase*).

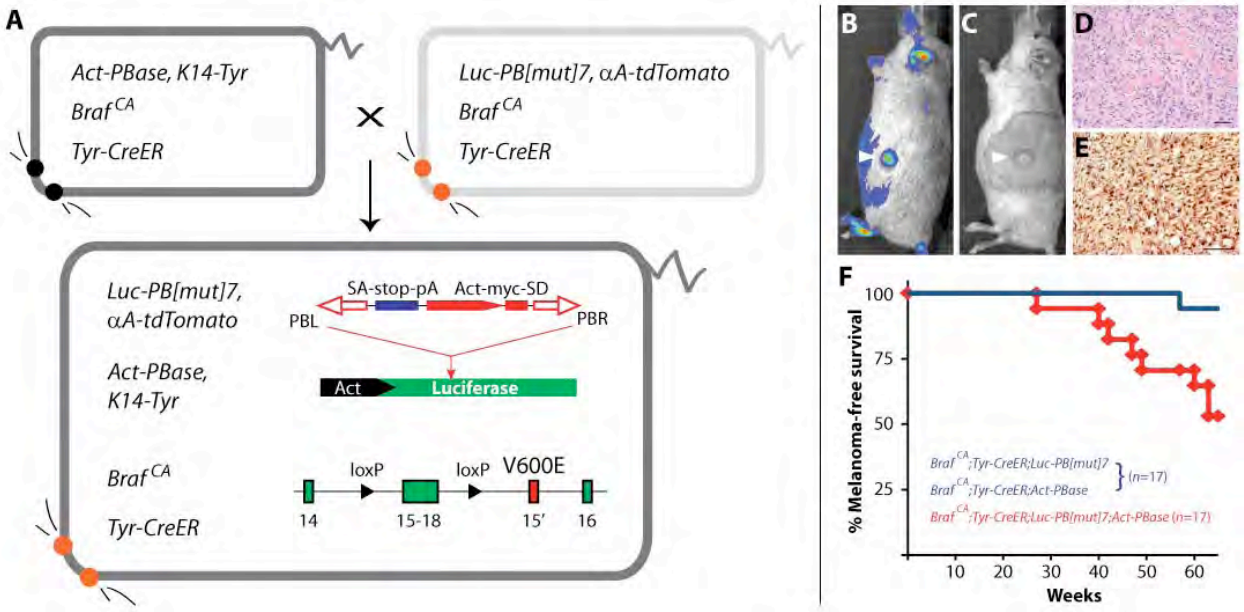

**Fig. 2.** Characterizing the nature of somatic insertional mutations. (**A**) Insertions into promoter regions of genes can force gene transcription, as in this one found in *Mitf*. (**B**) Exon-interrupting insertions like this one found in *Cdkn2a* result in reduced mRNA expression. (**C**) Two examples of how intron insertions (in either orientation) can disrupt gene expression. (**D**) Coding-direction insertions in introns can overexpress chimeric transcripts as in the case of chimeric *PB-Map3k2* exon 6 mRNA. Rel., relative transcript levels in inserted melanomas were compared to at least three uninserted melanomas and one representative experiment is shown (A-D).

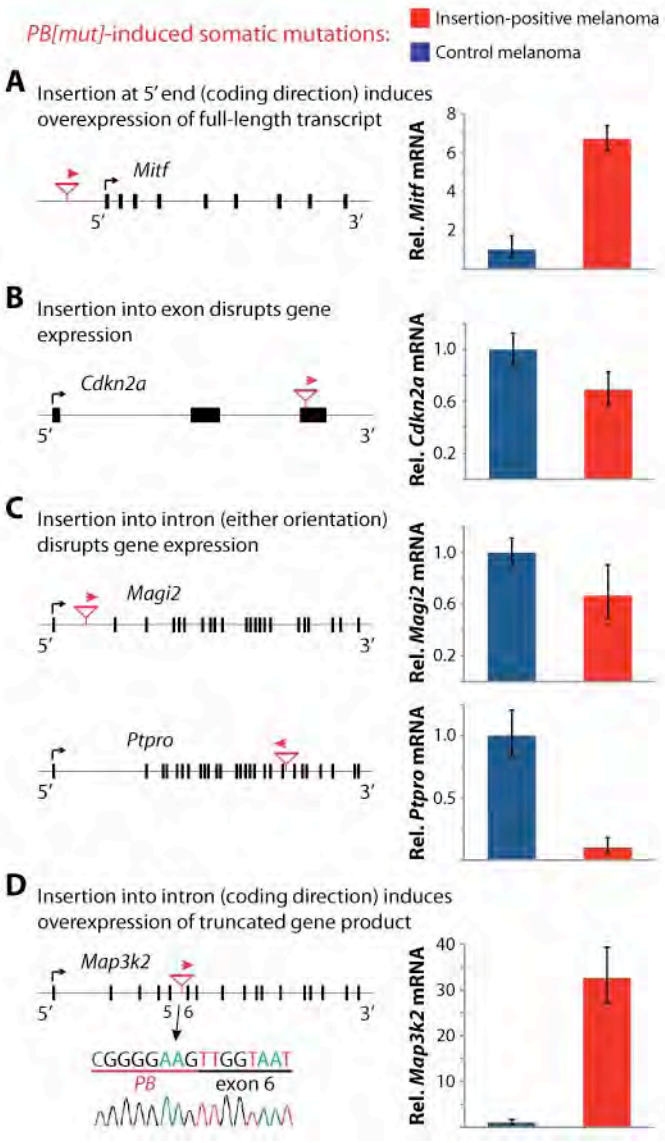

**Fig. 3.** A convergence of insertional mutations affecting *Map3k1* in multiple melanomas. (**A**) Insertions in *Map3k1* are clustered in introns 9 and 10, oriented in the coding direction. (**B**) Chimeric *PB-Map3k1* exon 10 and 11 mRNA can be detected by RT-PCR. (**C**) Truncated *Map3k1* transcripts are highly abundant in inserted tumors (red) versus controls (blue). Rel., relative transcript levels in inserted melanomas were compared to three uninserted melanomas and one representative experiment is shown. (**D** and **E**) Truncated Map3k1 proteins are highly expressed by melanomas harboring *Map3k1* insertions (D) but not by control uninserted melanomas (E, scale bar 25 µm). (**F**) PCR genotyping results reveal *Map3k1/2* insertional mutations are mutually exclusive with $Braf^{V600E}$ in the C30 and H12 melanomas. (**G**) Insertions affecting *Map3k1/2* and *Dusp4* in 45% of *PB*-induced melanomas (red) and mutations in *MAP2K1/2* for 7% of human melanomas (blue) (*7, 9*) suggest BRAF-independent ways to activate MAPK signaling promote melanomas.

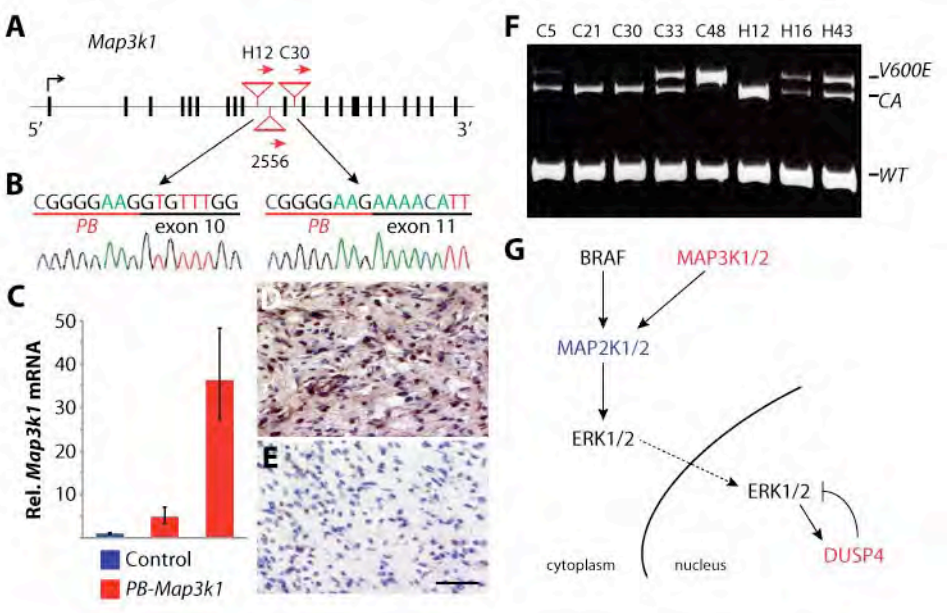

**Fig. 4.** Disruption of either *MAGI2* or *PTPRO* promotes anchorage-independent growth in human melanocytes. (**A**) *MAGI2* and genes encoding MAGI2 interactors are frequently mutated by *PB* insertions (red) and non-silent somatic mutations (blue numbers) in melanomas as seen in this STRING diagram (*30*). (**B**-**D**) Knock down of *MAGI2* and a randomly chosen gene from our group, *PTPRO*, as determined by mRNA (B and D) and protein levels (C) in immortalized human melanocytes stably expressing $BRAF^{V600E}$. Cells expressing a short hairpin with no homology to the mammalian genome (*shScramble*) were used as controls. Rel., relative transcript or protein levels. (**E**) Disruption of either *MAGI2* or *PTPRO* caused soft agar colony formation. 10X- and 50X-magnifications of representative soft agar colonies are shown (scale bar 200 µm).

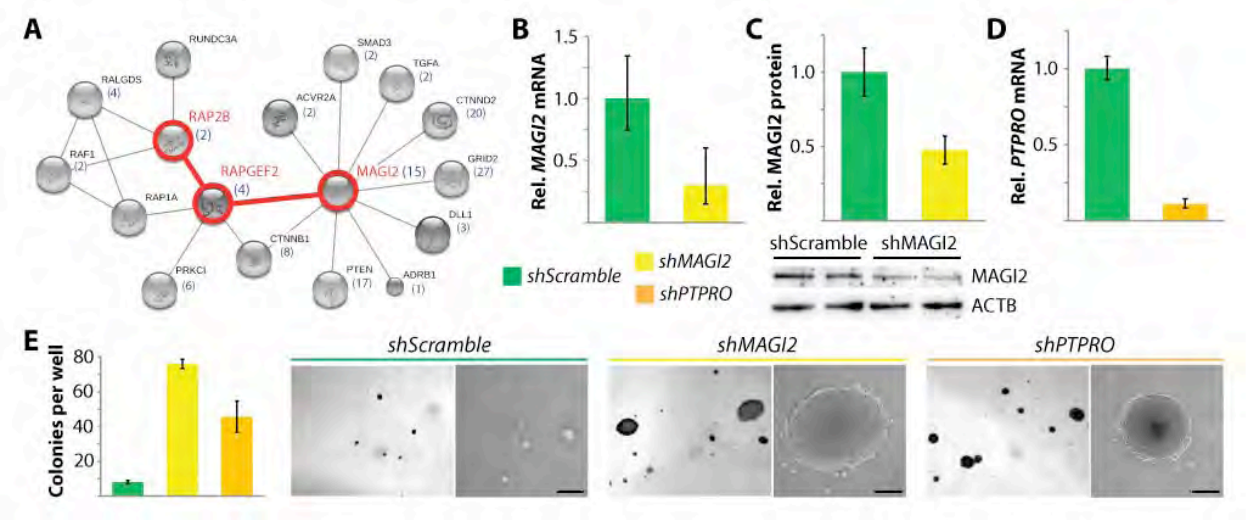

**Fig. 5.** Low-copy *PB* mutagenesis in mice can differentiate genes important for human melanomas. (**A**) Orthologs of 34 *PB*-inserted genes have mutations in human melanoma studies (ellipses and yellow circle) (*5-9, 29*) including known melanoma drivers (black). Four genes are mutated in 5/7 data sets, and nine genes are mutated in 4/7 data sets. (**B**) 28/37 inserted genes are mutated in $n \geq 3$ human melanomas (black). 2 genes are mutated in two human samples (dark gray). 4 genes are mutated in one sample (light gray). (**C**) Mutation rates in low-copy *PB* mutagenesis (red) are 100-fold lower than those in human melanomas (blue) (*5-9*). (**D**) For each gene identified in our screen, the observed number of mutations (x-axis) in human melanomas is plotted against the expected number of mutations (y-axis). Because of the high human melanoma mutation rate (orange line), only *CDKN2A* is significantly mutated. However, were the mutation rate ten-fold lower as in other cancers, genes mutated at these frequencies would have been identified as significantly mutated (green line).

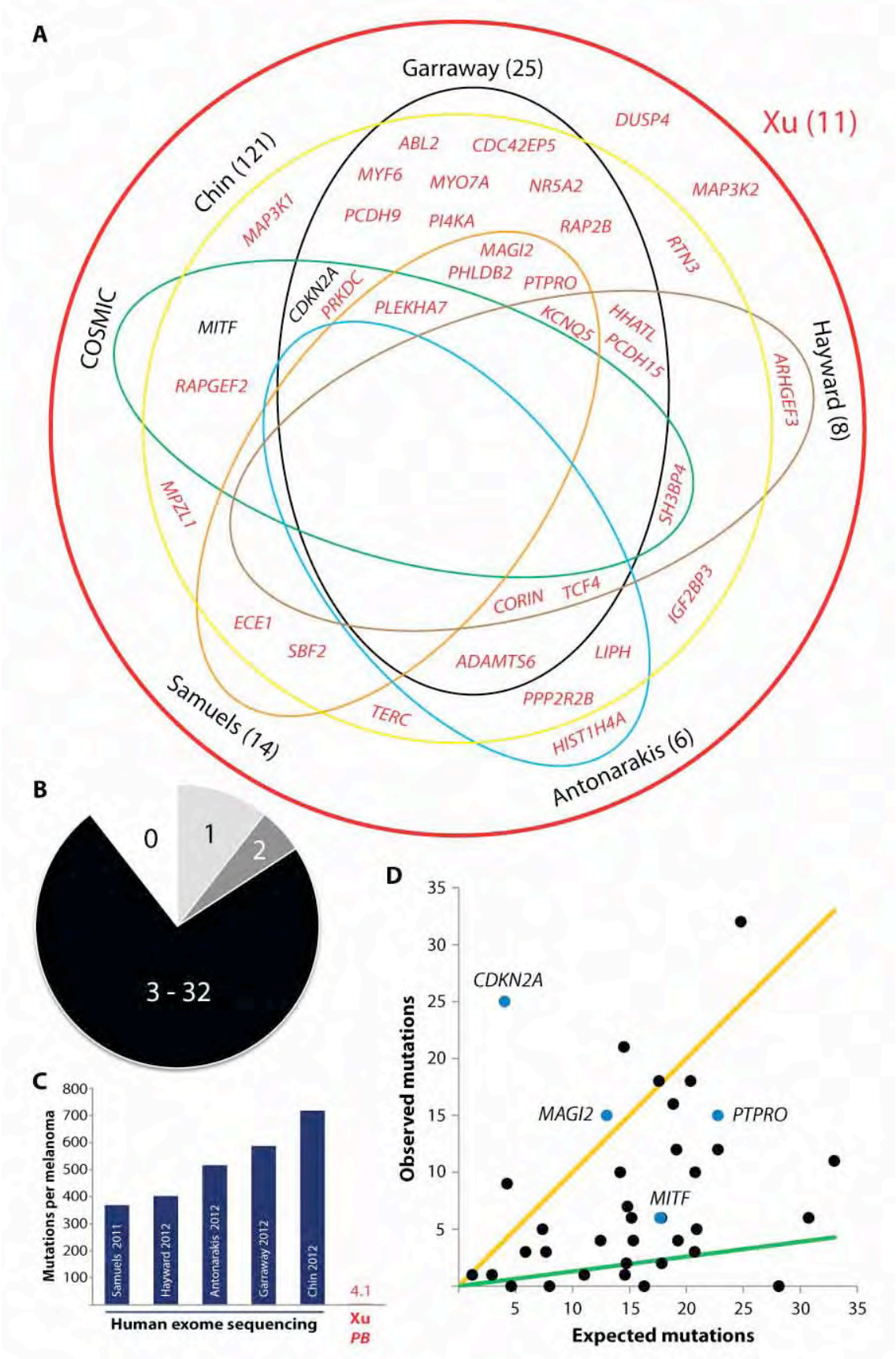
A
Garraway (25)
Xu (11)
Chin (121)
COSMIC
Hayward (8)
Samuels (14)
Antonarakis (6)
DUSP4
ABL2
CDC42EP5
MAP3K2
MYF6
MYO7A
NR5A2
MAP3K1
PCDH9
PI4KA
RAP2B
RTN3
MAGI2
PHLDB2
PTPRO
CDKN2A
PRKDC
PLEKHA7
KCNQ5
HHATL
MITF
PCDH15
ARHGEF3
RAPGEF2
SH3BP4
MPZL1
CORIN
TCF4
IGF2BP3
ECE1
SBF2
ADAMTS6
LIPH
PPP2R2B
TERC
HIST1H4A
B
0
1
2
3 - 32
C
Mutations per melanoma
800
700
600
500
400
300
200
100
Samuels 2011
Hayward 2012
Antonarakis 2012
Garraway 2012
Chin 2012
4.1
Xu
PB
Human exome sequencing
D
Observed mutations
Expected mutations
CDKN2A
MAGI2
PTPRO
MITF
35
30
25
20
15
10
5
5
10
15
20
25
30
35

**Table 1.** Recurrently mutated genes in low-copy *PB*-induced melanomas are shown here with the corresponding tumor name and *Braf* mutational status. Melanoma, melanoma sample name; *Braf* status, *Braf* mutational status (WT or V600E or mix of both).

| **Gene** | **Melanoma** | ***Braf* status** |
|---|---|---|
| *Arhgef3* | C48 | V600E |
| *Arhgef3* | C139 | WT |
| *Arhgef3* | H43 | mix of both |
| *Map3k1* | 2556 | WT |
| *Map3k1* | C30 | WT |
| *Map3k1* | H12 | WT |
| *Map3k2* | C30 | WT |
| *Map3k2* | C33 | mix of both |
| *Mitf* | 2556 | WT |
| *Mitf* | 3084 | WT |
| *Rapgef2* | C5 | WT |
| *Rapgef2* | H43 | mix of both |

**References and Notes:**


1. S. Nik-Zainal *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979-993 (2012).
2. The Cancer Genome Atlas Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-615 (2011).
3. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330-337 (2012).
4. E.D. Pleasance *et al*. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463**, 191-196 (2010).
5. X. Wei *et al*. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nat. Genet.* **43**, 442-446 (2011).
6. M. Stark *et al*. Frequent somatic mutations in MAP3K5 and MAP3K9 in metastatic melanoma identified by exome sequencing. *Nat. Genet.* **44**, 165-169 (2012).
7. S.I. Nikolaev *et al*. Exome sequencing identifies recurrent somatic MAP2K1 and MAP2K2 mutations in melanoma. *Nat. Genet.* **44**, 133-139 (2012).
8. M.F. Berger *et al.* Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature* **485**, 502-566 (2012).
9. E. Hodis *et al.* A landscape of driver mutations in melanoma. *Cell* **150**, 251-263 (2012).
10. Z. Ivics, P.B. Hackett, R.H. Plasterk, Z. Izsvak. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**: 501-510 (1997).
11. S. Ding *et al.* Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122**, 473-483 (2005).
12. N.G. Copeland, N.A. Jenkins. Harnessing transposons for cancer gene discovery. *Nat Rev Cancer* **10**, 696-706 (2010).
13. R. Rad *et al.* PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice. *Science* **330**, 1104-1107 (2010).
14. S.F. Landrette, J.C. Cornett, T.K. Ni, M.W. Bosenberg, T. Xu. piggyBac transposon somatic mutagenesis with an activated reporter and tracker (PB-SMART) for genetic screens in mice. *PLoS ONE* **6**, e26650 (2011), doi:10.1371/journal.pone.0026650.
15. H.N. March *et al.* Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. *Nat Genet.* **43**, 1202-1209 (2011).
16. K.M. Mann *et al.* Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc Natl Acad Sci U S A.* **109**, 5934-5941 (2012).
17. P.A. Perez-Mancera *et al.* The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. *Nature* **486**, 266-270 (2012).
18. H. Davies *et al.* Mutations of the BRAF gene in human cancer. *Nature* **417**, 949-954 (2002).
19. P.M. Pollock *et al.* High frequency of BRAF mutations in nevi. *Nat Genet.* **33**, 19-20 (2003).

20. D. Dankort *et al.* A new mouse model to explore the initiation, progression, and therapy of $BRAF^{V600E}$-induced lung tumors. *Genes Dev.* **21**, 379-384 (2007).
21. D. Dankort *et al.* $Braf^{V600E}$ cooperates with Pten loss to induce metastatic melanoma. *Nat Genet.* **41**, 544-552 (2009).
22. M. Bosenberg *et al.* Characterization of melanocyte-specific inducible Cre recombinase transgenic mice. *Genesis* **44**, 262-267 (2006).
23. B.T. Brett *et al.* Novel molecular and computational methods improve the accuracy of insertion site analysis in Sleeping Beauty-induced tumors. *PLoS ONE* **6**, e24668 (2011), doi:10.1371/journal.pone.0024668.
24. Information on materials and methods are available on Science Online.
25. L.A. Garraway *et al.* Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117-122 (2005).
26. J.C. Cronin *et al*. Frequent mutations in the MITF pathway in melanoma. *Pigment Cell Melanoma Res.* **22**, 435-444 (2009).
27. J.L. Blank, P. Gerwins, E.M. Elliott, S. Sather, G.L. Johnson. Molecular cloning of mitogen-activated protein/ERK kinase kinases (MEKK) 2 and 3. *J Biol Chem.* **271**, 5361-5368 (1996).
28. M. Karandikar, S. Xu, M.H. Cobb. MEKK1 binds Raf-1 and the Erk2 cascade components. *J Biol Chem.* **275**, 40120-40127 (2000).
29. S. Bamford *et al*. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br. J. Cancer* **91**, 355-358 (2004).
30. D. Szklarczyk *et al.* The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* **39**, D561-D568 (2011).
31. W.J. Kent. BLAT – the BLAST-like alignment tool. *Genome Res.* **12**, 656-664 (2002).
32. D. M. Bryant *et al.* A molecular network for de novo generation of the apical surface and lumen. *Nat Cell Biol.* **12**, 1035-1045 (2010).

**Acknowledgements:** We thank M. McMahon for $Braf^{CA}$ mice; J. Overton for Illumina sequencing; H. Widlund for p'mel-$BRAF^{V600E}$ cells; P. De Camilli, R. Halaban and A. Bacchiocchi for equipment, reagents and technical support; R. Lifton for critical discussion of the results. T.X. and T.K.N. designed the experiments, analyzed the data and wrote the manuscript. T.K.N. performed the experiments, S.F.L. contributed to the screen, R.D.B. contributed to mapping insertions, M.W.B. contributed mouse lines and performed histological analysis. This work was supported by the Howard Hughes Medical Institute.

**Supplementary Materials**

Materials and Methods

Figure S1

Tables S1-S3

References (*31-32*)

**Supplementary Materials:**

**Materials and Methods**

**Ethics statement.** Under protocol number 2008-10230, all animal experiments were approved by and conducted in compliance with the Yale Animal Resources Center and Institutional Animal Care and Use Committee.

**Mouse husbandry, genotyping and imaging.** *Act-PBase* and *Luc-PB[mut]7* mice were genotyped by coat color and fluorescent eye marker, respectively (*11, 13*). $Braf^{CA}$ and *Tyr-CreER* mice were genotyped as described previously (*20, 22*). To activate CreER, newly weaned mice were administered 1 mg tamoxifen mixture in sunflower seed oil for five consecutive days by intraperitoneal injection. For luciferase imaging, mice were sedated by isoflurane and injected with luciferin as described previously (*13*). To assess *Braf* allele recombination by PCR, $Braf^{V600E}$ (334 bp) and $Braf^{CA}$ (307 bp) products were resolved by 15% native PAGE.

**Histology and immunohistochemistry.** Melanomas were fixed in 10% neutral-buffered formalin overnight, dehydrated via increasing concentrations of ethanol and xylenes, paraffin-embedded and sectioned. Sections were stained with hematoxylin and eosin. For immunohistochemistry, heat-induced antigen retrieval in sodium citrate was performed on

unstained sections after rehydration. MAP3K1 (rabbit anti-MEKK1, Millipore #07-1476) and S100 (rabbit anti-S100 Ab-2, Thermo Scientific #RB-044-A1) antibodies were used. Slides were developed using a diaminobenzidine-peroxidase substrate kit (Vector Labs) and counterstained with hematoxylin.

**Mapping transposon insertion sites.** For initial examination of the nature of *PB* insertional mutations, we cloned and sequenced PCR-amplified melanoma insertion products by Sanger sequencing using a linker-mediated PCR protocol described previously (*13*). Results were mapped against the mouse genome using the BLAT tool on the UCSC Genome Browser (http://genome.ucsc.edu) (*31*). For Illumina high-throughput sequencing and mapping, indexed primers containing unique 8-bp tumor identifier sequences were used for the secondary PCR. Sequencing primers were designed to the outermost transposon ends (table S3). To determine potential driver insertional mutations, highly prevalent insertional mutations from Illumina sequencing reads were deemed likely to be present in the dominant tumor clone and drive tumorigenesis. Insertional mutations likely to be present in the dominant tumor clone were defined to have sequencing reads equal to at least 5% of the most abundant insertion. For each melanoma, insertional mutations of the dominant tumor clone accounted for 80-99% of all Illumina sequence reads. For verifying the prevalence of insertions in tumors, PCR was performed on genomic DNA corresponding to 15 insertional mutations using primers specific to the *PB* transposon and DNA flanking insertion sites (table S3).

**RT-PCR.** Melanomas were macrodissected and snap frozen in liquid nitrogen. Frozen tissue was homogenized using a bead mill homogenizer (Qiagen). Cell lysates were homogenized by passing through a 20-gauge syringe. Purified RNA was used with a blend of

oligo(dT), random primers and iScript reverse transcriptase (Biorad) to synthesize cDNA. Quantitative RT-PCR with SYBR green in the StepOne Real-Time PCR Instrument (Applied Biosystems) was performed in triplicate for each sample to determine mRNA levels using the relative standard curve method (table S3). For detection of chimeric mRNA, conventional RT-PCR was performed using primer pairs that span splicing junctions between the *PB[mut]* initiator exon and downstream exons of inserted genes (table S3). DNA bands were purified (Qiagen) and sequenced.

**Lentivirus infection and soft agar assay.** Short hairpin RNA specific for *MAGI2* and *PTPRO* (table S3) were cloned into *pLKO.1-blast* (Addgene 26655) (*32*). Lentiviruses were produced by overnight triple co-transfection of HEK293T cells in 10-cm dishes using Fugene HD and 4 μg *Δ8.9*, 2 μg *VSV-G*, and 4 μg of *pLKO.1-blast-shMAGI2*, *pLKO.1-blast-shPTPRO* or *pLKO.1-blast-shScramble* (Addgene 26701). Human melanocytes stably expressing $BRAF^{V600E}$ (p'mel-$BRAF^{V600E}$) (*25*) were incubated with 48- and 72-h post-infection viral supernatants supplemented with polybrene. Cells were selected with blasticidin for one week, after which they were seeded in 6-well plates at $10^4$ cells per well in 0.4% SeaPlaque agarose in DMEM/F12 with 10% FBS (Invitrogen) and layered onto 0.8% agarose in DMEM/F12 with 10% FBS. Soft agar experiments were performed in triplicate. Colonies were imaged, stained with 0.005% crystal violet and quantified (colonies ≥ 200 μm diameter) after three weeks.

**Western blot analysis.** Cells were lysed in NP-40 lysis buffer with protease and phosphatase inhibitors (Roche) and lysates were recovered by centrifugation. Total protein was quantified by the BCA Assay (Pierce). MAGI2 (rabbit anti-MAGI2 H-60, Santa Cruz Biotechnology sc-25664) and ACTB (mouse anti-β-Actin 8H10D10, Cell Signaling Technology 3700) antibodies

were used. Quantification of protein levels was performed by densitometric analysis of immunoblots using the ImageJ program (National Institutes of Health).

**Statistical Analysis.** The expected random mutational frequencies for genes being mutated by chance were calculated factoring in the size of the coding sequence for each gene and the average human melanoma mutation rate of 650 mutations per tumor (*5-9*). All calculations were performed under the assumptions that the probability of mutation at all possible positions of the exome is equal, each gene is mutated once per tumor and mutational events are independent. $\chi^2$ tests were used to assess the statistical significance of observed versus expected mutational frequencies. The Yates correction was applied as necessary.

**Analysis of protein interaction networks.** We utilized the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) to build networks of experimentally validated, functional human protein associations (*30*). A confidence score of 0.6 or higher was used.

**Fig. S1.** Verification of insertional mutations in melanomas by PCR across mapped genomic loci.

Transposon insertions in all 15 randomly chosen insertional mutations were detected by PCR.

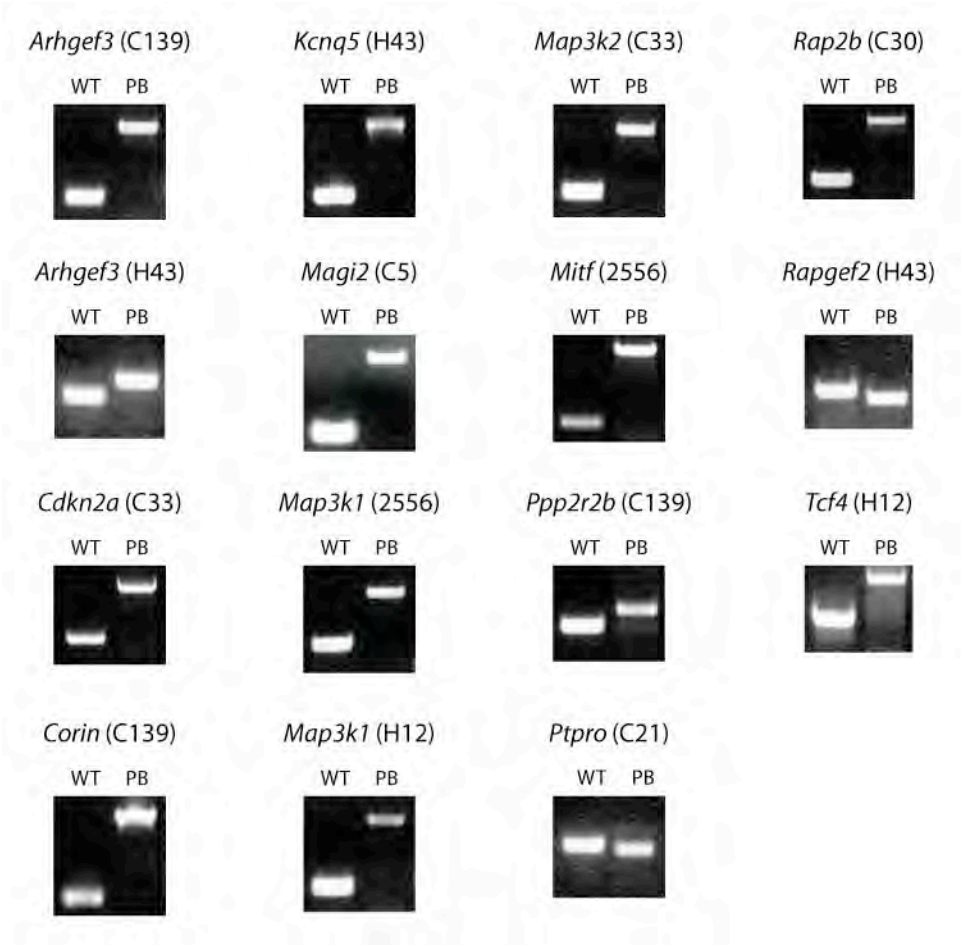

**Table S1.** Insertional mutations in 11 *PB*-induced melanomas.

| **Melanoma** | **Gene** | **Orientation relative** to **transcription** | **Insertion site in transcription unit** |
|---|---|---|---|
| 2556 | *Mitf* | Coding | 5' reg |
| 2556 | *Map3k1* | Coding | Intron 9 |
| 3084 | *Mitf* | Coding | 5' reg |
| 3084 | *Cdc42ep5* | Opposite | Exon 1 |
| 3084 | *Myo7a* | Coding | Intron 16 |
| C139 | *Corin* | Coding | Intron 1 |
| C139 | *Ppp2r2b* | Opposite | Intron 1 |
| C139 | *Arhgef3* | Coding | Intron 2 |
| C21 | *Plekha7* | Coding | 5' reg |
| C21 | *Ptpro* | Opposite | Intron 22 |
| C30 | *Rap2b* | Coding | 5' reg |
| C30 | *Terc* | Opposite | 5' reg |
| C30 | *Hhatl* | Opposite | Intron 1 |
| C30 | *Map3k1* | Coding | Intron 10 |
| C30 | *Ece1* | Coding | Intron 2 |
| C30 | *Igf2bp3* | Opposite | Intron 5 |
| C30 | *Map3k2* | Coding | Intron 5 |
| C30 | *Nr5a2* | Opposite | Intron 6 |
| C33 | *Cdkn2a* | Coding | Exon 3 |
| C33 | *Pcdh9* | Opposite | Intron 3 |
| C33 | *Map3k2* | Coding | Intron 5 |
| C48 | *Mpzl1* | Coding | Intron 1 |
| C48 | *Rtn3* | Coding | Intron 1 |
| C48 | *Arhgef3* | Coding | Intron 2 |
| C5 | *Dusp4* | Coding | 5' reg |
| C5 | *Rapgef2* | Coding | Intron 1 |
| C5 | *Magi2* | Coding | Intron 1 |
| H12 | *Pi4ka* | Opposite | Intron 15 |
| H12 | *Lypla1* | Opposite | Intron 2 |
| H12 | *Sh3bp4* | Coding | Intron 2 |
| H12 | *Myf6* | Opposite | Intron 2 |
| H12 | *Tcf4* | Coding | Intron 7 |
| H12 | *Map3k1* | Coding | Intron 9 |
| H16 | *Hist1h4a* | Coding | 5' reg |
| H16 | *Adamts6* | Coding | Intron 1 |
| H16 | *Phldb2* | Coding | Intron 1 |
| H16 | *Sbf2* | Coding | Intron 27 |
| H16 | *Pkig* | Coding | Intron 4 |
| H16 | *Pcdh15* | Coding | Intron 8 |
| H43 | *Kcnq5* | Opposite | Exon 15 |

| | | | |
|---|---|---|---|
| H43 | *Liph* | Opposite | Exon 2 |
| H43 | *Abl2* | Coding | Intron 1 |
| H43 | *Rapgef2* | Coding | Intron 1 |
| H43 | *Arhgef3* | Coding | Intron 2 |
| H43 | *Prkdc* | Coding | Intron 25 |

**Table S2.** Low-copy *PB* somatic mutagenesis identifies 33 genes also mutated in human melanoma.

| Gene | Melanomas mutated | References |
|---|---|---|
| *PCDH15* | 32 | (*6, 8, 9*) |
| *CDKN2A* | 25 | (*8, 9, 29*) |
| *CORIN* | 21 | (*6-9*) |
| *PLEKHA7* | 18 | (*5, 8, 9, 29*) |
| *PHLDB2* | 18 | (*5, 8, 9*) |
| *ADAMTS6* | 18 | (*7-9*) |
| *MAGI2* | 16 | (*5, 8, 9*) |
| *KCNQ5* | 16 | (*5, 6, 8, 9*) |
| *PTPRO* | 15 | (*5, 8, 9*) |
| *PRKDC* | 13 | (*8, 9, 29*) |
| *PI4KA* | 12 | (*8, 9*) |
| *TCF4* | 12 | (*6-9*) |
| *MYO7A* | 10 | (*8, 9*) |
| *MYF6* | 9 | (*8, 9*) |
| *NR5A2* | 7 | (*8, 9*) |
| *SH3BP4* | 7 | (*6, 9, 29*) |
| *MITF* | 6 | (*9, 29*) |
| *PCDH9* | 6 | (*8, 9*) |
| *ABL2* | 6 | (*8, 9*) |
| *PPP2R2B* | 5 | (*7, 9*) |
| *RAPGEF2* | 5 | (*9, 29*) |
| *MAP3K1* | 5 | (*9*) |
| *IGF2BP3* | 4 | (*9*) |
| *ECE1* | 4 | (*5, 9*) |
| *HHATL* | 3 | (*6, 8, 9*) |
| *LIPH* | 3 | (*7, 9*) |
| *SBF2* | 3 | (*5, 9*) |
| *MPZL1* | 2 | (*9*) |
| *RAP2B* | 2 | (*8, 9*) |
| *HIST1H4A* | 1 | (*7*) |
| *CDC42EP5* | 1 | (*8, 9*) |
| *ARHGEF3* | 1 | (*6*) |

| | | |
|---|---|---|
| *RTN3* | 1 | (*9*) |

**Table S3.** Primers and oligonucleotides used in this study.

| Name | Sequence | Notes |
|---|---|---|
| **For secondary LM-PCR amplification and multiplexed Illumina sequencing of insertion products** | | |
| SP2-8BP | 5'-CAAGCAGAAGACGGCATACGAGAT$N_8$AGTCAGTCA GCCGTGGCTGAATGAGACTGGTGTCGAC-3' | Linker-specific, 8-nt barcodes denoted by $N_8$ |
| IPBLSP2 | 5'-AATGATACGGCGACCACCGAGATCTACAC$N_6$GAGAG AGCAATATTTCAAGAATGCATGCGT-3' | PBL-specific, six-nt barcodes denoted by $N_6$ |
| IPBRSP2 | 5'-AATGATACGGCGACCACCGAGATCTACAC$N_6$ACGCA TGATTATCTTTAACGTACGTCACAA-3' | PBR-specific, six-nt barcodes denoted by $N_6$ |
| PBL-RD1 | 5'-TCAATTTTACGCAGACTATCTTTCTAGGGTTAA-3' | Read 1 primer (PBL) |
| PBR-RD1 | 5'-CGTACGTCACAATATGATTATCTTTCTAGGGTTAA-3' | Read 1 primer (PBR) |
| INDEX | 5'-CATTCAGCCACGGCTGACTGACT-3' | Index read primer |
| RD2 | 5'-GAGACTGGTGTCGACTCCGCTTAAGGGAC-3' | Read 2 primer |
| **For verification of insertional mutations in melanomas by genomic PCR** | | |
| LR1 | 5’-AAATCAGTGACACTTACCGCATTGACAAG-3’ | Sequence in PBL |
| LR2 | 5’-CCTCTAATAGTCCTCTGTGGCAAGGTCAAG-3’ | Sequence in PBL |
| 9M3k1-F | 5’-CAGGTGGGACCAGAAGAGAA-3’ | *Map3k1* |
| 9M3k1-R | 5’-CACCCCAGAATTGGGTTAAA-3’ | *Map3k1* |
| 10M3k1-F | 5’-CCGGCACAGTGAGTGTTCTA-3’ | *Map3k1* |
| 10M3k1-R | 5’-TTCTGGAGGTCAAACCCAAG-3’ | *Map3k1* |
| Arhg3-F1 | 5’-TCACTGGTGCACAATTAGGG-3’ | *Arhgef3* |
| Arhg3-R1 | 5’-TCCAGCAATGTGAAGCTGAC-3’ | *Arhgef3* |
| Arhg3-F2 | 5’-TGGGGGTTATTTGTGGAATG-3’ | *Arhgef3* |
| Arhg3-R2 | 5’-AGCTCGCGTTAGCACAACTT-3’ | *Arhgef3* |
| Cdkn2a-F | 5’-GCCACCTAGAGAGCCTAGCA-3’ | *Cdkn2a* |
| Cdkn2a-R | 5’-TTCTCATGCCATTCCTTTCC-3’ | *Cdkn2a* |
| Corin-F1 | 5’-GCGCCTGCATTTATGTGTAA-3’ | *Corin* |
| Corin-R2 | 5’-TTTGCTTTGTGACACCTTGG-3’ | *Corin* |
| Kcnq5-F | 5’-TCCGTTGGAGGGTACATCAT-3’ | *Kcnq5* |
| Kcnq5-R | 5’-GGGGGAGTGTGTCTGAAAAA-3’ | *Kcnq5* |
| M3k2-F | 5’-TGGACGAAGCAGATAGCAGA-3’ | *Map3k2* |
| M3k2-R | 5’-TCAGCTTGAACATTGCTTCAGT-3’ | *Map3k2* |
| Magi2-F | 5'-AACAGCTCCTAACATGAAAATGC-3' | *Magi2* |
| Magi2-R | 5'-AAATGGACCTTGATTTTTCCAA-3' | *Magi2* |
| Mitf-F | 5’-TCATAGTCCGTTTGCCATAAGC-3’ | *Mitf* |
| Mitf-R | 5’-ACATGCAGCCGTGTCTCTTT-3’ | *Mitf* |
| Ppp2r2b-F | 5’-ATGTGATCTCCCCCAGTGAG-3’ | *Ppp2r2b* |
| Ppp2r2b-R | 5’-GGATTCCCACGACTCCTACA-3’ | *Ppp2r2b* |
| Ptpro-F | 5’-TGGGAAAGAGCGAGACTGAT-3’ | *Ptpro* |

| | | |
|---|---|---|
| Ptpro-R | 5’-ACCGTAAGCCATCCAATCTG-3’ | *Ptpro* |
| Rap2b-F | 5’-CCAAACAGAGCTTTGGGAAG-3’ | *Rap2b* |
| Rap2b-R | 5’-TCTGAAGCTGAGAGGGGAAA-3’ | *Rap2b* |
| Rapgef2-F | 5’-TGTCTGCTTTATTCAATACATCAGCT-3’ | *Rapgef2* |
| Rapgef2-R | 5’-AGTACTGTGTGCCCTCTCCT-3’ | *Rapgef2* |
| Tcf4-F | 5’-GCATTTTGGTGTCAGCGACT-3’ | *Tcf4* |
| Tcf4-R | 5’-AGGAGGCTCAGCATACTGGA-3’ | *Tcf4* |
| **For detection of chimeric mRNA** | | |
| ME3-F | 5'-TGGGGAATGGAGCAGAAGCTC-3' | Transposon exon |
| M3k1-10R | 5'-ATAAGCAGCCAACGAGTTCC-3' | *Map3k1* exon 10 |
| M3k1-11R | 5'-TACCAGCATGGCTCTCAATG-3' | *Map3k1* exon 11 |
| M3k2-6R | 5'-CCACAGCTTTGTCCAAGTCA-3' | *Map3k2* exon 6 |
| **For quantitative RT-PCR** | | |
| Actbfw | 5'-CCCTAAGGCCAACCGTGAA-3' | *Actb* |
| Actbrv | 5'-CAGCCTGGATGGCTACGTACATG-3' | *Actb* |
| Cdkn2a-F | 5'-CGACGGGCATAGCTTCAG-3' | *Cdkn2a* |
| Cdkn2a-R | 5'-CTGAGGCCGGATTTAGCTCT-3' | *Cdkn2a* |
| Gapdh-6 | 5'-GGCATTGCTCTCAATGACAA-3' | *Gapdh* |
| Gapdh-7 | 5'-CCCTGTTGCTGTAGCCGTAT-3' | *Gapdh* |
| Magi2-F | 5'-CACAGGATGGGAAAATGGAG-3' | *Magi2* |
| Magi2-R | 5'-CTGCATGAGTGTGTCCAAGG-3' | *Magi2* |
| MAGI2-F | 5'-GTGAGGCTGCTGCTCAAGA-3' | human *MAGI2* |
| MAGI2-R | 5'-CCAGGGAGACGCCTACTTC-3' | human *MAGI2* |
| M3k1-17 | 5’-TCTCCACGAGAACCAGATCA-3’ | *Map3k1* |
| M3k1-18 | 5’-GCAATTCTCAGCCTCTGACC-3’ | *Map3k1* |
| M3k2-16 | 5’-TCAGTGGAGAAGGCTATGGAA-3’ | *Map3k2* |
| M3k2-17 | 5’-CTTCAAATTCAGCCCAAGGT-3’ | *Map3k2* |
| Mitf-F | 5'-GGCACCTGCTGCTCAGAGT-3' | *Mitf* |
| Mitf-R | 5'-GGGAAAGTCCATGCGCTCTA-3' | *Mitf* |
| Ptpro-21 | 5'-ATCTACCGCTGAACCGATGT-3' | *Ptpro* |
| Ptpro-22 | 5'-TCCTTCCTCTTCGTTCATGG-3' | *Ptpro* |
| Ptpro-24 | 5’-TTGTGTTCACAGTGCGACAG-3’ | *Ptpro* |
| Ptpro-25 | 5’-AGGGCAATGAAGGTTCCTG-3’ | *Ptpro* |
| PTPRO-F | 5'-AAGCAGCAGTTCTGCATCAGT-3' | human *PTPRO* |
| PTPRO-R | 5'-ATCTGGAAGCAAGGGAGGAT-3' | human *PTPRO* |
| **For knockdown of MAGI2 and PTPRO by shRNA** | | |
| MAGI2-t | 5'-CCGGCAAGCTGAACTTATGACCTTACTCGAGTAAGGT CATAAGTTCAGCTTGTTTTT-3' | |
| MAGI2-b | 5'-AATTAAAAACAAGCTGAACTTATGACCTTACTCGAGTA AGGTCATAAGTTCAGCTTG-3' | |
| PTPRO-t | 5'-CCGGGTTGCTTGTTACCCTCATTATCTCGAGATAATGA GGGTAACAAGCAACTTTTT-3' | |
| PTPRO-b | 5'-AATTAAAAAGTTGCTTGTTACCCTCATTATCTCGAGATA | |

| | | |
|---|---|---|
| | ATGAGGGTAACAAGCAAC-3' | |

PLoS one

# piggyBac Transposon Somatic Mutagenesis with an Activated Reporter and Tracker (PB-SMART) for Genetic Screens in Mice

**Sean F. Landrette[1☯], Jonathan C. Cornett[1☯], Thomas K. Ni[1☯], Marcus W. Bosenberg[2], Tian Xu[1,3]***

**1** Department of Genetics, Yale University School of Medicine, Boyer Center for Molecular Medicine, Howard Hughes Medical Institute, New Haven, Connecticut, United States of America, **2** Departments of Dermatology and Pathology, Yale University School of Medicine, New Haven, Connecticut, United States of America, **3** School of Life Science, Fudan-Yale Center for Biomedical Research, Institute of Developmental Biology and Molecular Medicine, Fudan University, Shanghai, China

## Abstract

Somatic forward genetic screens have the power to interrogate thousands of genes in a single animal. Retroviral and transposon mutagenesis systems in mice have been designed and deployed in somatic tissues for surveying hematopoietic and solid tumor formation. In the context of cancer, the ability to visually mark mutant cells would present tremendous advantages for identifying tumor formation, monitoring tumor growth over time, and tracking tumor infiltrations and metastases into wild-type tissues. Furthermore, locating mutant clones is a prerequisite for screening and analyzing most other somatic phenotypes. For this purpose, we developed a system using the *piggyBac* (PB) transposon for somatic mutagenesis with an activated reporter and tracker, called PB-SMART. The PB-SMART mouse genetic screening system can simultaneously induce somatic mutations and mark mutated cells using bioluminescence or fluorescence. The marking of mutant cells enable analyses that are not possible with current somatic mutagenesis systems, such as tracking cell proliferation and tumor growth, detecting tumor cell infiltrations, and reporting tissue mutagenesis levels by a simple *ex vivo* visual readout. We demonstrate that PB-SMART is highly mutagenic, capable of tumor induction with low copy transposons, which facilitates the mapping and identification of causative insertions. We further integrated a conditional transposase with the PB-SMART system, permitting tissue-specific mutagenesis with a single cross to any available Cre line. Targeting the germline, the system could also be used to conduct F1 screens. With these features, PB-SMART provides an integrated platform for individual investigators to harness the power of somatic mutagenesis and phenotypic screens to decipher the genetic basis of mammalian biology and disease.

**Citation:** Landrette SF, Cornett JC, Ni TK, Bosenberg MW, Xu T (2011) piggyBac Transposon Somatic Mutagenesis with an Activated Reporter and Tracker (PB-SMART) for Genetic Screens in Mice. PLoS ONE 6(10): e26650. doi:10.1371/journal.pone.0026650

**Editor:** Debra L. Silver, Duke University Medical Center, United States of America

**Received** July 18, 2011; **Accepted** September 29, 2011; **Published** October 21, 2011



**Funding:** This work was supported by the Howard Hughes Medical Institute as well as the following sources: JCC was supported by a fellowship from the National Cancer Institute (F32CA132320); SFL was supported by an Anna Fuller Fund Fellowship in Molecular Oncology and an American Cancer Society New England Division Fellowship; TKN was supported by the Genetics Training Grant (T32 GM007499) and the Training Program in Genomics and Proteomics Technologies (T32 HG003198) from the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* E-mail: tian.xu@yale.edu

☯ These authors contributed equally to this work.

## Introduction

Transposon insertional mutagenesis (TIM) is a powerful tool for inducing and identifying mutations of interest and has been utilized with great effect in many organisms, from the bacterium to the fruit fly *Drosophila melanogaster* [1,2]. The *Sleeping Beauty* (SB) and *piggyBac* (PB) DNA transposons have been recently developed for TIM in mice and human cells [3,4,5]. As one of the first applications of TIM in mammals, SB and PB have been used to identify tumor-promoting genes for multiple tumor types in mice [6–13]. In addition, SB transposase has been activated by tissue-specific Cre expression for the successful induction of colorectal adenocarcinoma, hepatocellular carcinoma, and B-cell lymphoma [10,11,13]. The successful induction of tumors by SB and PB indicates that a more versatile somatic TIM system can be established for broad application.

Somatic genetic screening requires animal breeding and significant lengths of time for manifestation of phenotypes such as tumorigenesis. Thus, it represents substantial investments of cost and time for mammals. Although tumor induction by SB TIM has been successful in multiple tissues, tumors in a number of other tissues have yet to be found despite employing the same mutator transposons and whole-body mutagenesis approach. This highlights the importance of determining the feasibility of TIM for the tissue of interest. Although one could molecularly determine mutagenesis efficiency before the manifestation of phenotype, it is not desirable to sacrifice valuable experimental animals during the course of the screen. It would be ideal to have a method to easily determine whether TIM in a targeted tissue is feasible at the outset of the screen without sacrificing animals.

Moreover, while tumors are readily identifiable, pre-tumorous lesions and metastatic clones are difficult to locate without visible markers. Furthermore, marking mutant clones is a prerequisite for screening and analyzing most other somatic phenotypes. Visibly marking mutant somatic clones has been employed in *Drosophila*,

zebrafish, and mice and demonstrated to have tremendous utility in analyzing a variety of processes including cell proliferation and migration, neurobiology and other clonal behaviors *in vivo* [14–17]. Thus, a somatic TIM system that can be used to screen for phenotypes other than tumor formation must incorporate the ability to track mutagenized cells.

Somatic mutagenesis screening has the genetic power to conduct genome-wide interrogation with a small number of animals. A highly efficient somatic TIM system would allow genetic screens to be performed with low copy numbers of transposon, thus allowing easy identification of causative mutations. Finally, it is optimal that the genetic elements for the features described above to be integrated such that somatic TIM can be conducted with one generation of breeding.

The PB transposon system has been shown to be highly efficient in mammals [3]. Its distinct properties of precise excision [3,18], large payload capacity [3], and genome-wide insertion tendency [3,12] offer a unique opportunity to create a somatic TIM system with broad utility. Here, we describe a method for PB somatic mutagenesis with an activated reporter and tracker in mice (PB-SMART). In this system, mobilization of the PB mutator transposon activates luciferase for the *ex vivo* reporting of mutagenesis activity levels, providing feasibility assessment at the outset of a genetic screen. The activation of luciferase or a red fluorescent protein also featured in this system further enables visual tracking of mutated cells. The integration of the mutagenesis reporter, mutant cell tracker, and mutagenic transposon into a single transgene greatly simplifies genetic breeding. Further incorporation of Cre-inducible PB transposase (PBase) transgenes allows tissue-specific somatic TIM screens to be conducted with one genetic cross. Finally, we show that PB-SMART can induce somatic phenotypes, track clonal behavior of mutated cells, and allow easy identification of causative insertions. Thus, PB-SMART provides a powerful means for individual investigators to decipher the genetic basis of mammalian biology and disease.

## Results

### Generating a PB mutator transposon for multiple genomic contexts

PB has previously been found to integrate at high frequency near or within genes [3]. To efficiently induce somatic mutations, we generated a PB mutator transposon (Figure 1A, *PB[mut]*) to induce ectopic gene expression in multiple genomic contexts. The CMV early enhancer/chicken *β-actin* promoter was cloned between the PB transposon arms to induce overexpression of genes downstream of a transposon insertion site. A myc epitope in all three reading frames and a splice donor were placed after the Actin promoter such that transcription and translation initiation can incorporate the myc tag sequence prior to splicing into endogenous exons. The CMV enhancer can also act as a bidirectional enhancer to upregulate genes. Additionally, in the event that *PB[mut]* inserts into an intron, an N-terminally truncated product can be ectopically induced. Insertions into introns can also produce C-terminally truncated gene products initiated by the endogenous promoter. A splice acceptor followed by a sequence with stop codons in all three frames and a poly(A) signal was added for this purpose. Expression of truncated genes may produce dominant active and dominant negative proteins. Thus, by design, the *PB[mut]* transposon can induce the ectopic expression of full-length or truncated endogenous genes in many different genomic contexts (Figure 1A). Indeed, *PB[mut]* can induce ectopic full-length and truncated gene expression in PB-SMART screens (see below).

### PB somatic mutagenesis with an activated luciferase reporter

In order to develop a somatic mutagenesis system that can induce mutations while simultaneously and non-invasively (*ex vivo*) reporting mutagenesis levels in tissues, we engineered a strategy to bioluminescently mark cells in which transposon mobilization has occurred. We took advantage of the highly sensitive luciferase reporter gene [19] by inserting *PB[mut]* at a TTAA tetranucleotide target site within the coding sequence (Figure 1A, *Luc-PB[mut]*), thereby preventing luciferase expression from an upstream Actin promoter. During mutagenesis, active PB transposase (PBase) catalyzes the precise excision of the PB mutator transposon and a full-length luciferase product is reconstituted (Figure 1A). Five *Luc-PB[mut]* transgenic lines were established and their ability to report PBase activity was tested by crossing to our previously described *Act-PBase* line [3]. The *Luc-PB[mut]7* line was identified as the most robust reporter for PB mutagenesis activity that can be easily assayed by *ex vivo* imaging (Figure 1B). *Luc-PB[mut]7* behaves as a faithful reporter for PBase activity, since it does not produce any luciferase signal in mice lacking PBase (Figure 1B).

To examine whether *Luc-PB[mut]7* can differentiate between varying levels of PBase activity, we generated transgenic mice expressing a PBase-Estrogen Receptor fusion transgene (*Act-PBaseER*). In the inactive state, the PBaseER protein is sequestered in the cytoplasm. Stimulation with tamoxifen shuttles PBaseER to the nucleus and activates its ability to drive PB transposition (Figure S1). *In vivo*, luciferase signal was not observed in newly weaned, untreated *Luc-PB[mut]7;Act-PBaseER* transgenic mice. However, transgenic *Act-PBaseER* can mobilize PB upon tamoxifen administration (Figure 1C) and luciferase signal can be detected one week after treatment, providing a tool for conditional PB TIM. Quantitative PCR measuring *PB[mut]* excision revealed that PBaseER activity levels were significantly lower than those of *Act-PBase* (Figure 1C). These results were corroborated by the luciferase mutagenesis reporter signal from *ex vivo* imaging (Figure 1C). Furthermore, we sometimes observed leaky luciferase signal in older, untreated mice housed together with tamoxifen-treated mice (data not shown). This signal, however, is at least an order of magnitude lower than signal from treated mice as measured by *ex vivo* imaging (data not shown). The ability to differentiate between differing levels of luciferase signal indicates that *Luc-PB[mut]7* is a sensitive reporter for PBase activity.

By real-time PCR analysis, the *Luc-PB[mut]7* transgenic line was found to contain seven copies of the mutagenic transposon. When crossed to *Act-PBase* mice, the number of live-born double transgenic pups was 93% of the expected number ($n>100$). Since mutagenesis in *Luc-PB[mut]7;Act-PBase* mice was observed in all tissues as reported by luciferase, we addressed whether this low copy line is sufficient to induce different phenotypes when mutagenesis is targeted in all tissues or specific organs. We found *Luc-PB[mut]7* can induce tumor formation as well as other phenotypes including cellular features (*e.g.*, cell infiltration and clonal expansion) and morphological defects (*e.g.*, alopecia) (also see below).

We found that body-wide PB-SMART can induce tumor formation in non-sensitized backgrounds. We observed that *Luc-PB[mut]7;Act-PBase* and *Luc-PB[mut]7;Act-PBase;Cre* mice developed tumors while conducting sensitized screens. Nine out of 32 of these *Luc-PB[mut]7;Act-PBase(Cre)* mice developed tumors within 50 weeks while only one out of seventeen *Act-PBase(Cre)* and zero out of 32 *Luc-PB[mut](Cre)* controls developed a tumor (Figure 1D). We have collected additional tumors from cohorts that have not reached 50 weeks and other non-sensitized cohorts containing floxed latent alleles of *Pten* loss or *Braf* activation [20,21]. Across

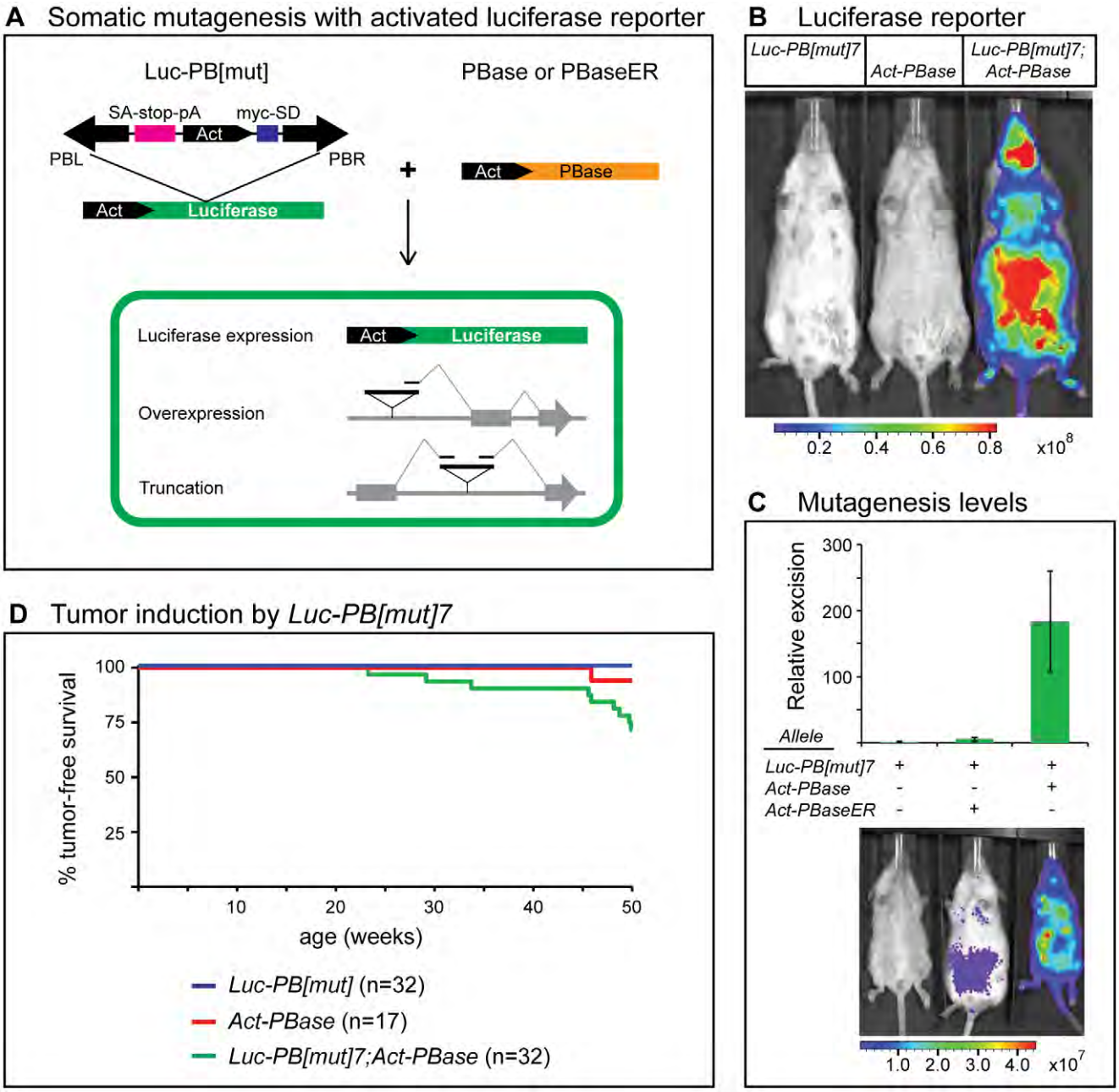


**Figure 1. Simultaneously inducing mutations and reporting mutagenesis activity levels with a PB transposon system.** (**A**) The Actin promoter (black pointed box) and myc-splice donor sequence (blue box) are engineered in the *PB[mut]* transposon to ectopically express downstream genes or partial transcripts. A splice acceptor and transcriptional termination sequences (pink box) truncate transcripts initiating upstream of the transposon insertion site. *PB[mut]* is inserted within an internal TTAA sequence of the luciferase gene to create *Luc-PB[mut]*. Transposase excises *PB[mut]* from the luciferase gene, restoring the luciferase expression and thus labeling cells. (**B**) By *ex vivo* luciferase imaging, PB mutagenesis occurs in *Luc-PB[mut]7;Act-PBase* mice (right) but not in single transgenic littermates (left and middle). Units, photons·s$^{-1}$·cm$^{-2}$·sr$^{-1}$. (**C**) Quantitative PCR analysis of transposon excision in the *Luc-PB[mut]7* line demonstrates a 30-fold increase in transposition activity between *Act-PBase* and *Act-PBaseER* (top). The difference in mutagenesis activity levels driven by *Act-PBaseER* (middle) and *Act-PBase* lines (right) can be differentiated by the luciferase PB reporter. Units, photons·s$^{-1}$·cm$^{-2}$·sr$^{-1}$. (**D**) Tumor-free survival curves over the span of 50 weeks show that *Luc-PB[mut]7;Act-PBase*(*Cre*) mice succumb to tumors while *Luc-PB[mut]7*(*Cre*) and *Act-PBase*(*Cre*) littermates remain relatively tumor-free.
doi:10.1371/journal.pone.0026650.g001

these cohorts, we have observed similar tumor types including kidney and pancreatic tumors, lung adenocarcinoma, hepatocellular carcinomas, soft tissue sarcomas such as angiosarcoma and small round blue cell (SRBC) tumors, and skin tumors such as squamous cell carcinoma, sebaceous cancer, and melanoma. These results confirm that our low copy PB-SMART system can efficiently induce tumor-promoting mutations in a wide variety of tissues.

We also found that the low copy mutagenic transposon in the PB-SMART system allows for easy identification of potential causative insertional mutations. Transposon insertions from tumors were mapped by linker-mediated PCR (LM-PCR). On

average, across 45 tumors analyzed, 6.1 PB insertions were mapped. Notably, greater than 64% of all insertions were recovered in introns, exons, or within ten kilobase pairs (kb) of either known or predicted genes (Table 1). Given the propensity of *Luc-PB[mut]* for inserting into or near genes in these tumors, we analyzed whether the mutator transposon was oriented along or opposite the coding direction of inserted genes. Because our mutator transposon is designed to overexpress full-length genes or truncated genes, we expected that in tumors, insertions oriented along the coding direction of nearby or affected genes would be enriched versus insertional orientation by random chance. Indeed, over 64% of intron insertions were oriented along the coding direction of the affected genes, and almost 60% of insertions within 10 kb upstream of genes were oriented toward the coding direction of the nearby genes (Table 2). Meanwhile, insertions within 10 kb downstream of nearby genes displayed less bias for mutator orientation (Table 2). These data imply that the PB mutator transposon is indeed upregulating genes important for tumorigenesis.

In five kidney tumors, we identified common insertions in *Mitf*. The mutator transposons in all five tumors were inserted in the coding direction, upstream of the translational start of the M-isoform, suggesting overexpression of *Mitf* in these tumors (Figure 2A). Indeed, quantitative PCR revealed that *Mitf* was upregulated in these kidney tumors as compared to two wild-type kidney samples (Figure 2B). Histopathology of the kidney tumors revealed an expansion of nests of pleomorphic spindle and epithelioid cells suggestive of carcinoma (Figure 2C).

Notably, two other members of the MiT family of basic-helix-loop-helix/leucine zipper transcription factors, *TFE3* and *TFEB*, have been recurrently affected by translocations that result in the upregulation of these transcription factors in human pediatric renal cell carcinoma [22–26]. Since TFE3, TFEB, and MITF are highly homologous and bind to the same DNA consensus sequence, it is likely that they can substitute for each other as oncogenes [27]. Together, these data suggest that *MITF* upregulation is an important driver for kidney tumor formation.

We also identified *Gli2*, a downstream effector of the *Hedgehog* (*Hh*) signaling pathway [28], as a common insertion site in SRBC tumors. We mapped coding-direction insertions in intron 7 or intron 8 from 11 tumors displaying SRBC morphology (Figures 2D and 2E). Gli2 contains a repressor domain in the amino-terminus, and expression of the C-terminal portion of Gli2 has previously been shown to result in the constitutive transcriptional activation of downstream target genes [29,30]. In these sarcomas, it is likely that the *Gli2* insertions lead to the expression of constitutive transcriptional activator forms of Gli2 that drive tumor formation. In fact, constitutive Hedgehog signaling in Gorlin's syndrome due to mutation in *PTCH* leads to an increased propensity to develop an SRBC tumor, rhabdomyosarcoma [31,32]. Furthermore, activation of Gli1 (a Gli2 transcriptional target) seems to be a major downstream effector of the EWS-FLI1 oncoprotein which drives another SRBC tumor, Ewing Sarcoma [33,34]. Together these data strongly support the notion that constitutive Hedgehog signaling is a critical driver in SRBC tumor formation, with overexpression of truncated *Gli2* being one mechanism. Thus, the identification of cancer genes with known human relevance validates the utility of our low-copy mutator for uncovering disease genes.

**Table 1.** Insertional frequency in tumors induced by *Luc-PB[mut]7* and *PB[mut-RFP]7*.

| PB insertions in | % of total hits |
|---|---|
| Genes (including ±10 kb) | 64.8 |
| known genes | 57.5 |
| predicted genes | 7.3 |
| Exons | 1.8 |
| Introns | 48.0 |
| 5′ regulatory sequences (≤10 kb) | 9.9 |
| 3′ regulatory sequences (≤10 kb) | 4.8 |

doi:10.1371/journal.pone.0026650.t001

**Table 2.** Mutator orientation in tumors induced by *Luc-PB[mut]7* and *PB[mut-RFP]7*.

| PB insertions in | % of hits in coding direction |
|---|---|
| Exons | 60.0 |
| Introns | 64.1 |
| 5′ regulatory sequences (≤10 kb) | 59.3 |
| 3′ regulatory sequences (≤10 kb) | 53.8 |

doi:10.1371/journal.pone.0026650.t002

## PB somatic mutagenesis with an activated RFP cell tracker

In somatic mutagenesis, the ability to track mutated cells via the same transgenic line used to induce mutations opens up many experimental possibilities. We thus added new elements in our PB mutator transposon to label the mutant cells. The RFP coding sequence and an internal ribosome entry site (IRES) were inserted between the Actin promoter and myc-SD, which contains start codons in all three reading frames followed by a splice donor, of *PB[mut]* (Figure 3A, *PB[mut-RFP]*). Following insertions into introns or 5′ regulatory regions, production of bicistronic pre-mRNA containing RFP, IRES, the engineered myc-SD exon, and downstream endogenous intron and exon sequences can be initiated by the Actin promoter. After splicing of the myc-SD exon to downstream endogenous splice acceptors, the bicistronic transcripts enables the expression of RFP and the ectopic expression of endogenous proteins (Figure 3A). When we tested this mutator in cultured cells, the activated RFP marker was co-expressed in cells that ectopically expressed downstream endogenous proteins that incorporated the engineered myc-SD exon (Figure 3B). In order to ensure that *PB[mut-RFP]* does not express RFP from a transgenic concatamer, a concatamer silencer (CS) was inserted outside of the transposon (Figure 3A). The CS contains a splice acceptor followed by stop codons in all three reading frames to terminate transcripts initiated by the Actin promoter from transgene concatamers. To prevent alternative splicing–mediated exclusion of the CS splice acceptor, a sequence forming a pre-tRNA-like (ΔAC) structure was cloned after the stop codons. This substrate is efficiently recognized and cleaved by RNase P ensuring that pre-mRNA which contain the CS do not form mature transcripts [35].

We generated seven transgenic lines varying in transposon copy number from two copies to 200 copies (Figure 3C) and found that none were RFP-positive, indicating that the CS successfully prevents undesired RFP production. When the two highest copy lines were crossed to the *Act-PBase* line, live-born double transgenic progeny were rarely observed (Figure 3D), suggesting that the mutagenic transposons can be mobilized by *Act-PBase* during embryonic development and can induce significant deleterious effects and embryonic lethality. The mobilization of the low copy

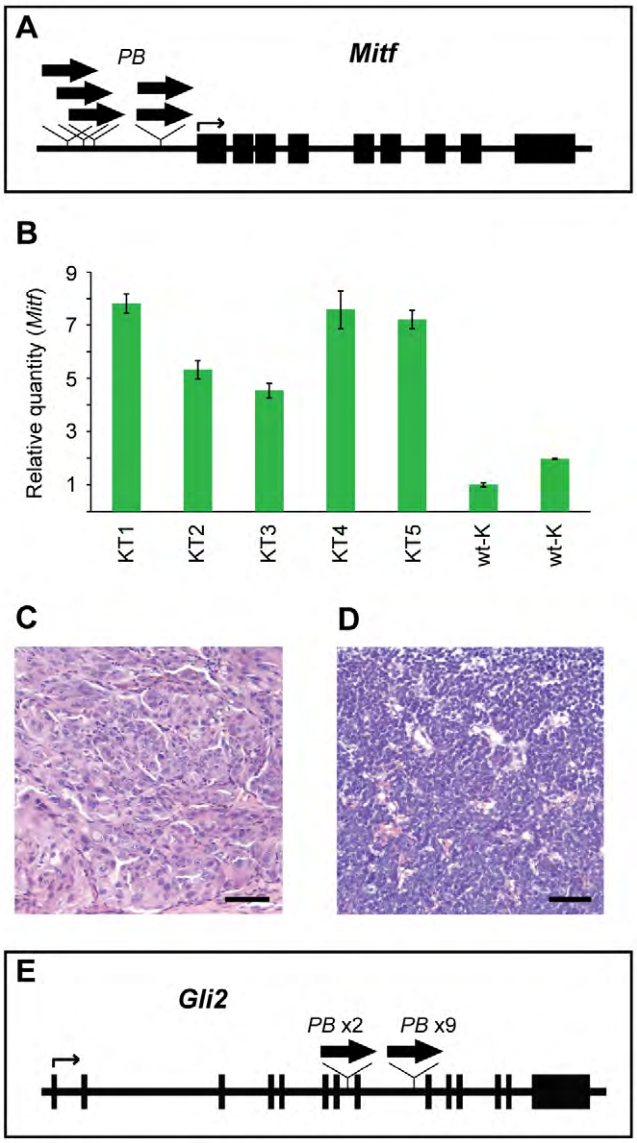


**Figure 2. Identification of driver oncogenes in solid tumors from *Luc-PB[mut]7;Act-PBase* mice.** (**A**) Insertions (wide arrows) upstream of the M-isoform of *Mitf* were mapped from five kidney tumors. (**B**) Quantitative PCR reveals that *Mitf* transcripts are upregulated in all five kidney tumors (KT1, KT2, KT3, KT4, and KT5) compared to two wild-type kidneys (wt-K). (**C**) Kidney tumors possessed similar histological profiles, featuring packets of spindle-shaped epitheliod cells, indicating carcinoma (scale bar, 50 μm). (**D**) Histological analysis shows SRBC morphology (scale bar, 100 μm). (**E**) Insertions (wide arrows) in intron 7 or 8 of *Gli2* were mapped from eleven SRBC tumors.
doi:10.1371/journal.pone.0026650.g002

*PB[mut-RFP]* lines was also confirmed by excision PCR (Figure 3E). Importantly, mutagenesis using either of the two-copy or the seven-copy *PB[mut-RFP]* lines crossed to *Act-PBase* mice resulted in the appearance of RFP-positive patches (Figure 3F) and triggered tumorigenesis (Figure 3G), validating these three low copy *PB[mut-RFP]* lines for somatic mutagenesis with an activated cell tracker.

## Tissue-specific PB somatic mutagenesis screening with an activated reporter and tracker (PB-SMART screen)

We have shown that the collective luciferase signal in tissues generated by *Luc-PB[mut]7* can act as a robust reporter of mutagenesis activity. Since *Luc-PB[mut]7* is also a mutator, the luciferase signal could also be used to mark the mutant cells. Given that *Luc-PB[mut]7* is a low copy line and not every cell in a tissue will be luciferase positive, we reasoned that clones or patches of mutant cells will have elevated signals and could be detected through *ex vivo* imaging. In this sense, the luciferase reporter also behaves as a tracker for mutant cells. By monitoring *ex vivo* luciferase signal, we identified clonal patches in the skin and tracked the appearance of melanomas in *Luc-PB[mut]7;Act-PBase;Braf$^{CA}$;Tyr-CreER* mice (Figures 4A and 6A; also see below). In addition, the luciferase signal enabled the identification of tumor cell infiltration into distant organs of a *Luc-PB[mut]7;Act-PBase;Pten$^{+/-}$* mouse that developed lymphoma (Figure 4B). The enlarged thymus and lymph nodes were strongly luciferase-positive due to the dense clustering of cells that activated the luciferase reporter (Figure 4B). Strong signal was also seen emanating from a distinct portion of the liver (Figure 4B). Histology revealed that the luminescent spot was indeed infiltration of lymphocytes (Figure 4C), and we mapped identical transposon insertions from the liver infiltration and other tumor sites, indicating that the cells originated from the same clone. Thus, the *Luc-PB[mut]7* line is a mutant clone tracker for monitoring clonal expansion, tumor formation, and infiltrations of cells into foreign tissues.

We further expanded the PB-SMART system for tissue-specific genetic screens. To allow the system to be readily adopted for screens in many types of tissues and cells, we generated transgenic mice in which *PBase* or *PBaseER* can be activated by tissue-specific Cre by inserting a Stop-pA sequence flanked by loxP sites between the Actin promoter and the transposase coding sequence (Figure 5A, *LSL-PBase*, *LSL-PBaseER*). When these double transgenics were crossed into lines expressing Cre in the kidney and genitourinary tract or the skin epithelium [36,37], correct targeting of mutagenesis to these tissues was confirmed by the luciferase reporter (Figures 5B and 5C).

To ask whether *Luc-PB[mut]7* can be used to track mutation-driven expansion of cells, we utilized the *Tyr-CreER* line, which expresses CreER specifically in melanocytes [38], one of the least densely clustered cell types in the body. *Tyr-CreER* has been used to conditionally activate *Braf$^{V600E}$*, thereby promoting melanocyte proliferation [20,21]. We generated *Luc-PB[mut]7;LSL-PBase;Braf$^{CA}$;Tyr-CreER* quadruple transgenic mice and assayed whether luciferase can track mutation-driven expansion of cells *ex vivo*. After administering tamoxifen, we were able to detect luciferase signal in melanocytes in the ears, tail, and dorsal trunk and also track the expansion of the positive cells over time (Figure 5D).

To further illustrate that *Luc-PB[mut]7* can be used to track clonal expansion in specific cell types, we targeted the PB-SMART system to the mouse skin. We generated *Luc-PB[mut]7;LSL-PBase;K14-Cre* mice which express Cre specifically in the basal keratinocytes of the epidermis. This compartment contains the stem and progenitor cells that give rise to the hair follicle, sebaceous gland, and interfollicular epidermis [39]. We reasoned that if a keratinocyte stem or progenitor cell were to be targeted by PB-SMART, over time an entire clonal unit would become luciferase positive. Indeed, the appearance of distinct luciferase clusters could be detected in these mice (Figure 6A).

Skin-specific PB-SMART resulted in tumor development as well as other morphological phenotypes. For example, we observed hair loss on dorsal skin, resulting from PB mutagenesis and marked by luciferase signal (Figure 6B). Keratin staining of frozen sections revealed a striking epidermal thickening in the affected region (Figures 6C and 6D). We mapped a transposon insertion upstream of *Myc* in the affected skin (Figure 6E). This is consistent with previous reports that in *K14-c-Myc* transgenic mice,

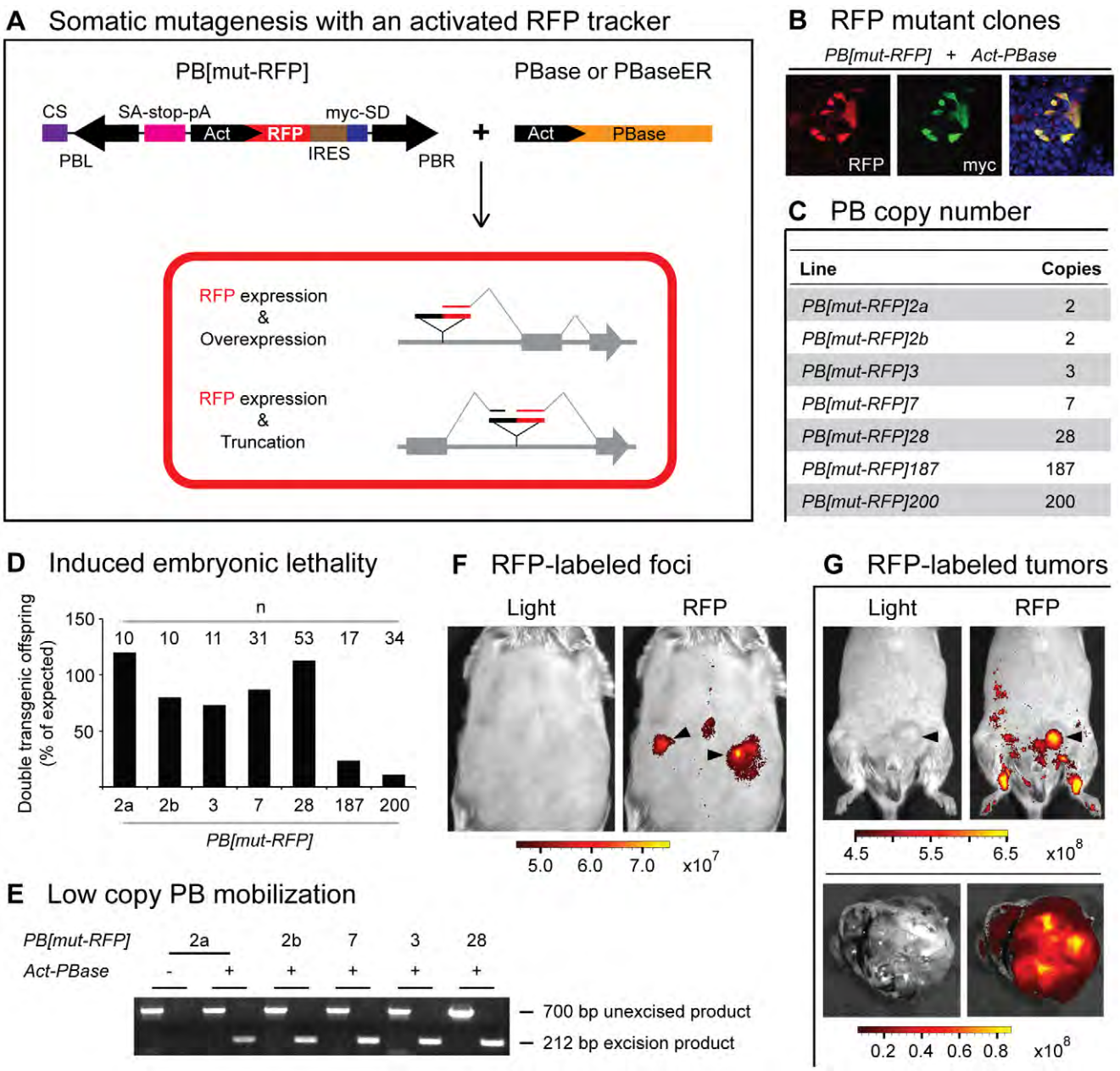


**Figure 3. Simultaneously inducing mutations and tracking mutated cells with a PB transposon system.** (**A**) *PB[mut-RFP]* couples RFP expression with ectopic expression of a downstream gene or partial gene transcript via the IRES (brown box). CS (purple box) prevents unwanted RFP expression from transgene concatamers. (**B**) *PB[mut-RFP]* mobilization in HEK 293 cells induces RFP-positive clones that colocalize with myc-tagged proteins ectopically expressed by the mutator transposon. (**C**) Transgenic lines carrying two to 200 copies of *PB[mut-RFP]* were generated. (**D**) Significant PB transposition-induced embryonic lethality was observed when high-copy lines were crossed to *Act-PBase*. (**E**) Low-copy *PB[mut-RFP]* lines crossed to *Act-PBase* mobilized the mutator transposon as determined by PCR with primers directed against sequences flanking the transposon. A 212-bp excision product was specifically detected in double transgenics but not in single transgenic littermates. A 700-bp unexcised product was observed in all *PB[mut-RFP]* mice. (**F**) Shown here in the dorsal skin of a *PB[mut-RFP]2a;Act-PBase* mouse, RFP-positive foci can be detected *ex vivo* over background signal. Units, photons·$s^{-1}$·$sr^{-1}$·$\mu M^{-1}$. (**G**) RFP-marked tumor formation (arrowhead) can be visualized *ex vivo* as seen in the upper panels showing a *PB[mut-RFP]2b;Act-PBase* mouse. In the lower panels, a stomach tumor dissected from a *PB[mut-RFP]7;Act-PBase* mouse is strongly RFP-positive. Units, photons·$s^{-1}$·$sr^{-1}$·$\mu M^{-1}$.
doi:10.1371/journal.pone.0026650.g003

epidermal stem cell proliferation and differentiation results in hair loss, hyperproliferation, and thickening of the interfollicular epidermis [40,41].

We further analyzed skin tumors that were induced and labeled by skin-specific PB-SMART. We were able to observe tumors in *Luc-PB[mut]7;LSL-PBase;K14-Cre* mice starting at four months of age. Using the e*x vivo* reporter system, we can detect regions of luciferase activity prior to tumor formation, that correspond to the later development of tumors (Figure 6F). In four of these tumors, we identified transposon insertions in *Gli2* (Figure 6G). Quantitative PCR confirmed that *Gli2* mRNA levels were increased in these tumors as compared to wild type skin (Figure 6H).

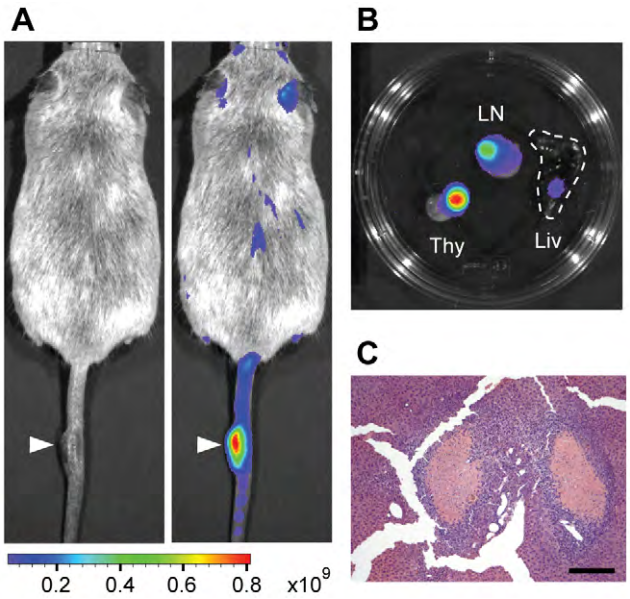


**Figure 4. Tracking tumor formation and infiltration with *Luc-PB[mut]7*.** (**A**) Luciferase signal labels a melanoma in this *Luc-PB[mut]7;Act-PBase;Braf$^{V600E}$;Tyr-CreER* mouse (white arrowheads). The left panel shows a light image while the right panel overlays the luciferase signal. (**B**) Luciferase signal enabled the identification of tumor cell infiltration into distant organs of a *Luc-PB[mut]7;Act-PBase;Pten$^{+/-}$* mouse that developed lymphoma. The enlarged thymus (Thy) and lymph nodes (LN) were strongly luciferase-positive. Strong signal was also seen emanating from a distinct portion of the liver (Liv, white dashed line). (**C**) H&E staining of liver from the luciferase-positive region reveals extensive lymphocyte infiltration (scale bar, 200 μm).
doi:10.1371/journal.pone.0026650.g004

Constitutive activation of the *Hh* pathway is characteristically seen in basal cell carcinoma (BCC) [42], which arises from stem cells in the epidermis and hair follicle [43,44]. Furthermore, histopathology of the skin tumors revealed palisading basaloid cells forming tumor nests and retraction of the stroma from the tumor islands in formalin-fixed sections similar to human BCC (Figure 6I). Together, these data show that the PB-SMART system can be deployed to conduct tissue-specific genetic screens and identify causative mutations relevant to human disease.

## Discussion

In the data presented, we have taken advantage of the PB transposon's properties to create a somatic mutagenesis system, PB-SMART, to simultaneously mutate and mark cells for conducting tissue-specific genetic screens in mice. In somatic screens for cancer genes, PB-SMART offers advantages over other TIM systems. Our system allows one to detect tumors before their morphological manifestation and to monitor tumor growth and dissemination *ex vivo*. Importantly, it allows one to detect tumor formation *ex vivo* such that only tumor-positive animals will be sacrificed in a screen. This is particularly critical for screens of internal cancers as previous screens required sacrificing large numbers of experimental animals at arbitrary time points to search for tumors. The visible cell markers also enable mutant clones in tissues to be readily identified and continuously monitored over time. This is essential for the detection of phenotypes other than tumors and provides many experimental advantages for visualizing proliferation defects, tracking cell migration, tracing cell lineages, and analyzing other clonal behavior *in vivo*. As proof-of-principle examples, we demonstrated that the system can be used to identify phenotypes in PB mutant clones, detect cell infiltration, and track the expansion of rare cell populations. In addition to these examples, the luciferase and RFP mutator lines can be utilized in other ways to confer specific experimental advantages according to the biological process to be surveyed. For example, luciferase is highly useful for detection and tracking of solid tumors originating from deep tissues. On the other hand, RFP enables detection and isolation of cells via fluorescence-activated cell sorting which would be particularly useful for capturing disseminated tumor cells in the bloodstream. As an additional resource to facilitate use by the research community, we will provide further detailed characterization of these and other recently developed PB-SMART lines on our lab website (http://info.med.yale.edu/genetics/xu).

The luciferase reporter integrated into the PB-SMART system also allows mutagenesis efficiency to be visually assessed at the onset of a genetic screen without sacrificing animals. Furthermore, we have found that luciferase signal activated by PB mutagenesis is highly indicative for the successful outcome of a mutagenesis screen, providing a valuable predictive tool. Genetic screens in mice are a significant investment. Since phenotypes in many screens take significant time to manifest, it is a tremendous advantage to have a predictive indicator at the beginning of the experiments.

PB-SMART is also a highly efficient mutagenesis system. Utilizing PB's large payload capacity, we engineered elements in PB to ectopically express genes or parts of genes and to terminate transcripts upon insertion in many genomic contexts, thereby creating highly mutagenic transposons. Together with PB's high transposition efficiency and genome-wide distribution, PB-SMART is able to induce somatic phenotypes with low copy (seven or fewer) transposons. The efficiency of the system is highlighted by the fact that a broad spectrum of tumors is induced with only seven copies of transposon. Importantly, low transposon copy numbers allow for more straightforward identification of causative insertions because there are fewer bystander insertions. We found that over 64% of insertions in tumors are mapped in introns, exons, or within 10 kb of genes. This is a remarkably high gene insertional frequency indicative that there is a selective advantage for PB gene insertions in tumors. In two tumor types that have not previously been found in somatic mutagenesis screens, we used common insertion site analysis to identify *Mitf* as a driver in kidney tumor formation and truncated *Gli2* as a driver of SRBC tumors. Thus, PB-SMART is a powerful tool to identify mutations relevant to human disease.

For the creation of a tissue-specific mutagenesis system that can be readily used by the research community for a variety of tissue and cell types, we generated conditional transposase transgenic lines that can be activated by Cre (*LSL-PBase* and *LSL-PBaseER*). We showed that mutagenesis can be targeted specifically to the skin or internal organs by crossing to tissue-specific Cre lines. Temporal control of tissue-specific mutagenesis can also be achieved by using *LSL-PBaseER* or by crossing to available CreER lines, as we demonstrated in melanocytes with *Tyr-CreER*. Given that specific Cre or CreER lines have been generated for almost every mouse tissue being studied, our system will make it possible to target any tissue for genetic screens. The combination of conditional transposase lines and low copy PB mutator lines could be used to target the germline for the generation of stable and heritable mutations for traditional screens. More importantly, mutations targeted at the germline could also be used to conduct

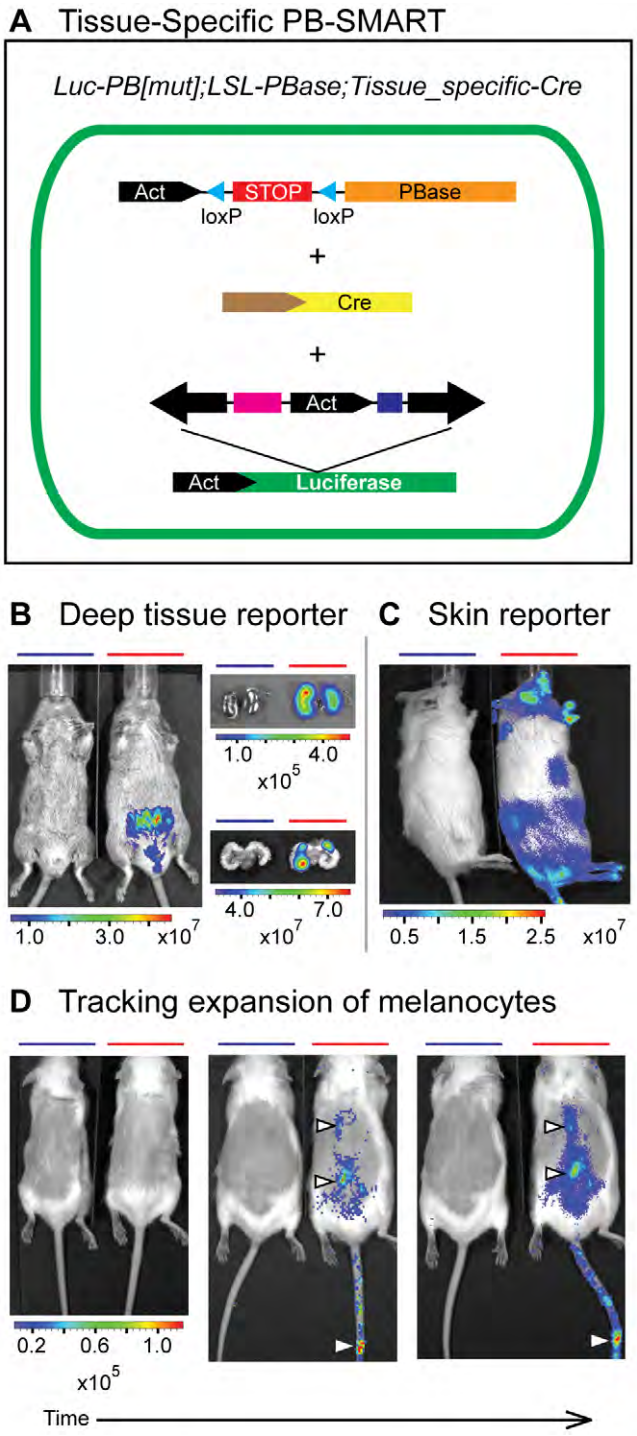


**Figure 5. Tissue-specific PB-SMART.** (**A**) Activation of the conditional *LSL-PBase* enables tissue-specific mutagenesis by crossing transgenic mice to available Cre lines. (**B**) *Luc-PB[mut]7;LSL-PBase;Ksp-Cre* transgenics (red bars) undergo PB mutagenesis specifically in the seminal vesicles and kidneys, reported by luciferase signal. Control mouse: *Luc-PB[mut]7; LSL-PBase* (blue bars). Units, photons·s$^{-1}$·cm$^{-2}$·sr$^{-1}$. (**C**) PB mutagenesis throughout the skin of *Luc-PB[mut]7;LSL-PBase;K14-Cre* mice (red bar) is reported by luciferase signal. Control mouse: *Luc-PB[mut]7;K14-Cre* (blue bar). Units, photons·s$^{-1}$·cm$^{-2}$·sr$^{-1}$. (**D**) *Tyr-CreER* conditionally activates $Braf^{V600E}$ and PB mutagenesis specifically in melanocytes of quadruple transgenic mice (*Luc-PB[mut]7;LSL-PBase;*$Braf^{CA}$*;Tyr-CreER*, red bars). The expansion of this cell population at three, nine, and eleven months (left, middle and right, respectively) is tracked *ex vivo* by luciferase signal. Control mouse: *Luc-PB[mut]7;BrafCA;Tyr-CreER* (blue bars). Units, photons·s$^{-1}$·cm$^{-2}$·sr$^{-1}$.
doi:10.1371/journal.pone.0026650.g005

F1 screens for developmental and physiological phenotypes including embryonic and adult defects.

In summary, the ability to induce somatic mutations, report mutagenesis levels, and track mutated cells permits many types of applications and analyses that are not possible with current TIM systems. Because induction and labeling of mutant cells are simultaneously achieved by the same mutator transposon, the design not only links the two events, but also simplifies genetics by minimizing the number of transgenes involved. Furthermore, PB-SMART also integrates the conditional transposase, which enables TIM screens to be conducted for any tissue in the mouse by a single genetic cross to a tissue-specific Cre line. The ability to screen thousands of clones in a single animal in only one generation makes it possible to achieve genome-wide interrogation with a small number of animals in a short period of time. The development of PB-SMART thus empowers individual investigators to decipher the genetic basis of biology and disease by forward genetic screens.

## Methods

### Ethics Statement

All experiments were approved by and conducted in compliance with the Yale Animal Resources Center and the Institutional Animal Care and Use Committee under protocol number 2008-10230.

### Plasmid construction

The *Luc-PB[mut]7* and *PB[mut-RFP]* transposons were generated from elements in the plasmids *PB[Act-RFP]* [3], *pIRES* (Clontech), *pBS302* (Life Technologies), *pΔAC* [35], and *pGL3* (Promega). Myc-SD was PCR amplified from an oligonucleotide containing ATG transcriptional initiation codons, myc epitope sequences in all three reading frames, and a splice donor sequence. CS and part of SA-stop-pA were PCR amplified from an oligonucleotide containing a splice acceptor and stop codons in all three reading frames (sequences available upon request). For *PB[mut-RFP]*, loxP sites were included to flank the mutagenic *Act-RFP-IRES-myc-SD* cassette. All elements were constructed on a *pBlueScript II SK+* (Stratagene) backbone before cloning between PBL and PBR transposon arms to yield the final plasmid. For *Luc-PB[mut]*, luciferase from *pGL3* was introduced outside PBL and PBR by overlapping PCR followed by cloning into the XhoI and NotI sites in *PB[Act-RFP]*.

*LSL-PBase* plasmid was generated from DNA elements in the plasmids *Act-PBase* [3] and *pBS302*. A 1.5-kb transcriptional termination sequence flanked by loxP sites was obtained from the *pBS302* plasmid by digestion with EcoRI and SpeI, and subsequently cloned between the Actin promoter and PBase coding sequence in the *Act-PBase* plasmid. Two copies of the chicken *β-globin* HS4 core enhancer from the plasmid *pNI-CD* were also cloned upstream of the Actin promoter.

*PBaseER* was generated by inserting a 12-nucleotide linker sequence that encodes a four-amino acid, positively charged linker (5′-AGGCGCGCCCCC-3′) between the *PBase* coding sequence and the mutated tamoxifen-responsive mouse estrogen receptor hormone-binding domain coding sequence [45]. *PBaseER* was subsequently cloned after the Actin or LSL promoter.

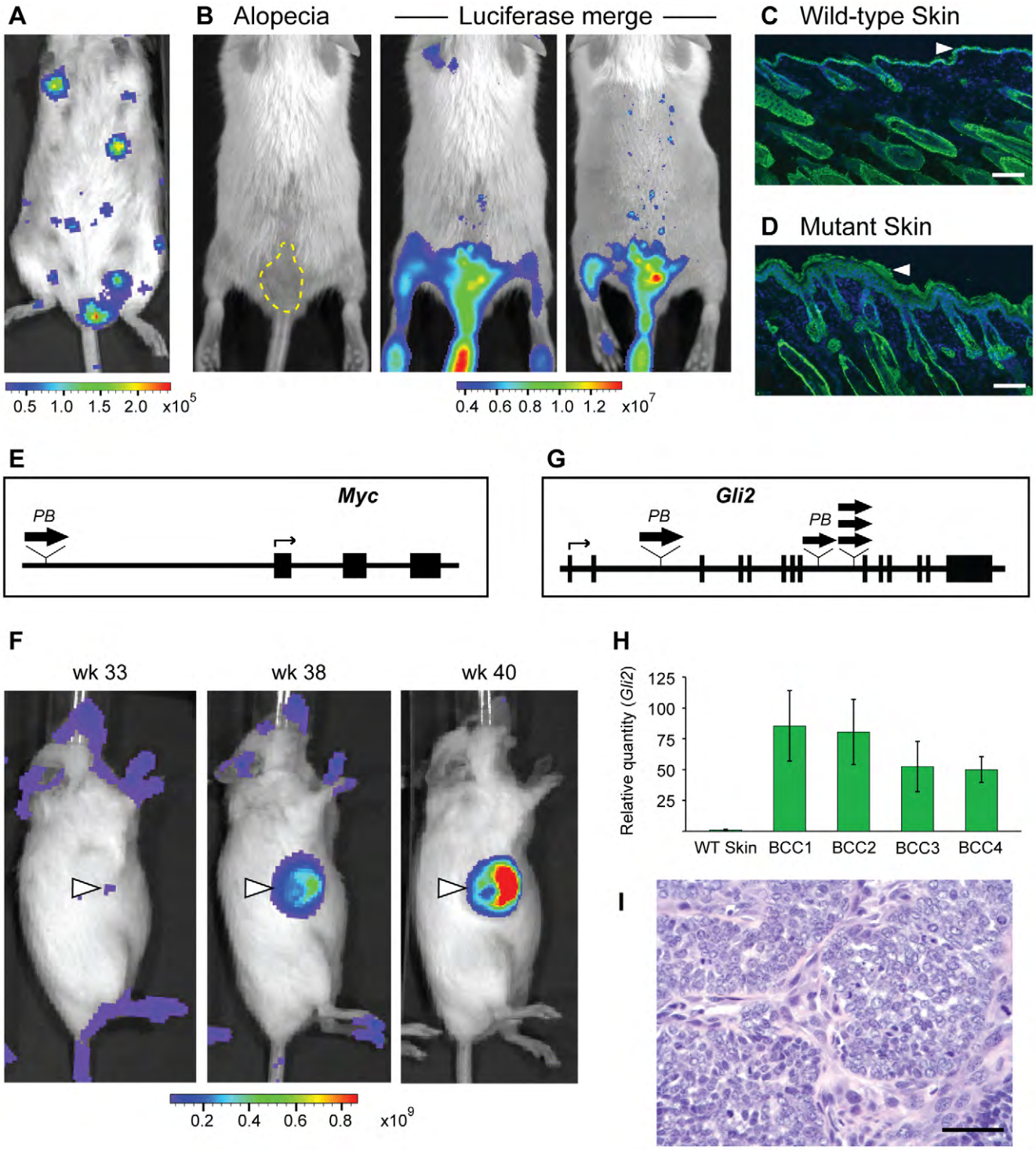


**Figure 6. Tissue-specific induction of skin phenotypes by Cre-activated PB mutagenesis.** (**A**) Luciferase signal marks clones of mutated cells in *Luc-PB[mut]7;LSL-PBase;K14-Cre* mice. (**B**) PB-induced alopecia in a *PB[mut]7;LSL-PBase;K14-Cre; Pten*$^{lox/+}$ mouse, (yellow dashed line) shows strong luciferase signal corresponding to hair loss (middle). Fully shaved, luciferase signal demarcates the specific region of alopecia (right). (**C**) Cytokeratin staining (green) of wild-type dorsal skin shows the normal thickness of the mouse epidermis (white arrowhead). DAPI (blue) marks cell nuclei. (**D**) Staining as in (c) from the region of alopecia reveals substantial thickening of the epidermis (white arrowhead). (**E**) A transposon insertion upstream of *Myc* was mapped to the affected region of skin. (**F**) Skin-specific mutagenesis in *Luc-PB[mut]7;LSL-PBase;K14-Cre* mice induces tumor formation and luciferase-labeled tumor cells. The white arrow corresponds to region of luciferase signal that develops into a larger tumor over time. Units, photons·s$^{-1}$·cm$^{-2}$·sr$^{-1}$. (**G**) Insertion sites (wide arrows) in *Gli2* were mapped in multiple skin tumors. (**H**) By quantitative PCR, *Gli2* transcripts were upregulated in each skin tumor (BCC1, BCC2, BCC3, and BCC4) compared to WT skin. (**I**) The skin tumors bear resemblance to human basal cell carcinoma according to histological analysis, consistent with the activation of the *Hedgehog* pathway (scale bar, 50 μm).
doi:10.1371/journal.pone.0026650.g006

## Generation of transgenic mice

*Luc-PB[mut]* and *PB[mut-RFP]* plasmids were linearized by NotI or PspXI. *LSL-PBase* plasmid was linearized by SalI and StuI. *Act-PBaseER* was linearized by SalI and BamHI. After gel purification (Qiagen), linear DNA was injected into C57BL/6J x SJL/J F2 hybrid one-cell embryos. Founders were screened for the presence or absence of the transgene by PCR analysis (see Table S1 for primer sequences).

## Mouse lines, breeding and husbandry

*Act-PBase*, *Tyr-CreER*, *Braf*$^{CA}$, *Pten*$^{lox/+}$, *Ksp1.3-Cre (Tg(Cdh16-cre)91Igr/J)*, and *K14-Cre (Tg(KRT14-cre)1Amc/J)* lines have been reported previously [3,20,36,37]. *Pten*$^{+/}$ were generated by mating a *Pten*$^{lox/+}$ male to a *Pbsn-Cre4 (B6.Cg-Tg(Pbsn-cre)4Prb)* female to induce recombination in the germline [46]. Transgenic mice were backcrossed to the FVB/NJ mouse strain. Animals were fed a standard rodent diet.

## Administration of tamoxifen to mice

100 mg/ml tamoxifen was dissolved in ethanol and diluted 1:10 in sunflower seed oil (for intraperitoneal injection) or 1:100 in water (for tamoxifen-water bottles). Newly weaned mice were injected with 100 µl (1 mg) tamoxifen-in-oil mixture once per day for five consecutive days. For tamoxifen-water bottles, mice were constantly kept on 100 µg/ml tamoxifen-water.

## Histology and immunofluorescence

For paraffin-embedded sections, tissues were dissected from euthanized mice and fixed in 10% neutral-buffered formalin overnight, dehydrated in a series of increasing concentrations of ethanol and xylenes, embedded in paraffin, sectioned, and stained with hematoxylin and eosin. For frozen sections, dissected tissues were OCT-embedded, frozen, sectioned, and fixed in 4% paraformaldehyde (PFA). For immunofluorescence staining of cytokeratin, OCT sections were fixed for 10 min in 4% PFA in PBS and washed three times for 5 min in PBS. Sections were permeabilized with 0.1% Triton for 15 min, washed three times for 5 min in PBS, and blocked for one hour with blocking solution (5% NGS, 1% BSA, and 0.1% Triton X-100 in PBS). Rabbit anti-cow cytokeratin wide spectrum screening antibody (1:250, Dako) was added overnight at 4°C. Alexa Fluor 488 goat anti-rabbit IgG (1:2000; Invitrogen) was used for detection of primary antibodies and nuclei were stained with DAPI. Blocking solution and antibody incubation steps were performed in a humidified chamber.

## *Ex vivo* mouse imaging

Mice were sedated using an isoflurane vaporizer. For *ex vivo* visualization of RFP, sedated *PB[mut-RFP]* control and *PB[mut-RFP];Act-PBase* mice were shaved and immediately imaged side-by-side in the IVIS Spectrum (Xenogen) by epifluorescence. Autofluorescence in shaved regions was quantitated in control mice to adjust epifluorescence imaging parameters. To visualize *in vivo* luciferase activity using the IVIS Spectrum, sedated mice were weighed, treated with 150 mg luciferin per kg body weight via intraperitoneal injection, and subsequently imaged 15 minutes after luciferin injection.. Heart rate and respiration were closely monitored and body temperature was regulated throughout the entire procedure.

## Linker-mediated PCR identification of PB insertion sites

LM-PCR mapping of PB mutator insertion sites was performed using a previously developed protocol [47] with the following modifications. Two µg of mouse genomic DNA were digested overnight with Sau3AI or MspI and subsequently heat inactivated. Double-stranded Sau3AI-specific or MspI-specific linkers were ligated to digested genomic DNA in an overnight reaction at 18°C. Ligation products were digested in a second reaction by MfeI to eliminate adaptor-ligated DNA fragments containing mutator transposon at its original site of insertion. Following heat inactivation of enzyme and purification of digestion products (Qiagen), a primary PCR reaction, followed by a nested PCR reaction, were performed. Amplification products were purified (Qiagen), TA-cloned into the *pGEM-T-Easy* vector (Promega), and transformed in competent cells. Colonies were screened for inserts by colony PCR using the T7 and SP6 primers and those containing inserts sequenced by traditional Sanger sequencing. Oligonucleotides used in this procedure can be found in Table S1.

## Sequence analysis and annotation

Insertion sites were mapped by Blatting against the mouse genome with UCSC Genome Bioinformatics Site (http://genome.ucsc.edu/).

## RNA isolation and quantitative PCR

RNA was isolated from tumors or wild type tissues (Qiagen). 1–2 µg of RNA was then used with oligo-dT primer and Superscript III (Invitrogen) to synthesize cDNA. Quantitative PCR with SYBR green in the StepOne Real-Time PCR instrument (Applied Biosystems) was used to determine transcript levels with the relative standard curve method. Primers are detailed in Table S1.

## Excision PCR

Quantitative PCR was used to determine the relative excision levels of *Luc-PB[mut]* from spleen genomic DNA. Mobilization of *PB[mut-RFP]* from the concatamer was determined by traditional PCR using primers designed against the plasmid sequences flanking the PB[mut-RFP] transposon. Primer sequences are detailed in Table S1.

## Cell culture

HEK 293 cells were grown in DMEM containing 10% FBS under standard conditions. Cells were transfected with circular plasmids using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). To select neomycin-resistant colonies, Geneticin (G418) was added to the culture medium at a concentration of 500 µg/ml. For two weeks, culture medium was replaced every 3–4 days and G418 was resupplemented each time. To induce PBaseER activity, 4-hydroxtamoxifen (4-OHT) was added to culture medium to obtain the desired final concentration.

## Confocal microscopy

HEK293 cells were seeded onto sterilized cover slips, cultured until 90% confluent, and then fixed in 4% paraformaldehyde (PFA) for 15 minutes at room temperature. Following permeabilization using 0.1% Triton X-100 in PBS (PBST), cells were blocked with 5% normal goat serum at room temperature for 30 minutes and then incubated with primary antibodies. The ERα primary antibody (sc-542, Santa Cruz Biotechnology) and mouse monoclonal anti-myc 9E10 primary antibodies were used. FITC-conjugated secondary antibodies were used to visualize proteins. Samples were mounted in Vectashield containing DAPI (Vector Laboratories). Immunofluorescence images were collected using a Zeiss LSM Meta confocal microscope and processed using Adobe Photoshop.

### Methylene blue staining

To stain colonies in tissue culture dishes, culture medium was aspirated by vacuum and washed once in PBS. Subsequently, cells were fixed in ice cold methanol for 20 minutes at room temperature. Staining was performed using 0.2% methylene blue solution for one hour at room temperature. Following this step, the culture dishes were rinsed with distilled water to remove residual stains.

### Western blotting

Cells were lysed in lysis buffer and protein concentrations of lysates were determined by the BCA Assay (Pierce). Approximately 15 μg of total protein was loaded on a 10% SDS-PAGE gel. Protein was transferred onto the Trans-Blot Transfer Medium (Biorad). Nitrocellulose membranes were blocked at room temperature for 1 hour with 5% milk in PBS, then washed 3 times for 10 minutes in PBS and incubated with ERα primary antibody (sc-542, Santa Cruz Biotechnology) in PBS and 3% BSA overnight at 4°C. Membranes were washed 3 times for 10 minutes in PBS, blocked at room temperature for 1 hour with 5% milk in PBS, and incubated in HRP-conjugated secondary antibodies. Immunoreactive bands were visualized using Western Lightning *Plus*-ECL (Perkin-Elmer).

## Supporting Information

**Figure S1 Tamoxifen-dependent translocation of PBaseER to the nucleus drives PB transposition in HEK 293 cells.** (**A**) *Act-PBaseER* in transfected HEK 293 cells was detected by Western blotting using a polyclonal ERα antibody and displayed the expected size. (**B**) Immunofluorescent staining with ERα antibody revealed that PBaseER translocates to the nucleus after 24-hour 4-hydroxytamoxifen (4-OHT) treatment. (**C**) HEK 293 cells co-transfected with circular plasmids carrying PB with a neomycin resistance cassette (*PB[SV40-neo]*) and *Act-PBase* form G418-resistant clones after drug selection for two weeks, indicative of PBase-mediated transposition activity. (**D**) 4-OHT treatment is required for PBaseER-mediated transposon insertion. Compared to *Act-PBase*, *Act-PBaseER* mediates transposition at a lower efficiency.
(TIF)

**Table S1** Primers used in this work for genotyping transgenic lines, analyzing PB excision, analyzing PB copy number, analyzing relative transcript levels, and mapping PB insertion sites.
(DOC)

## Acknowledgments

We thank M. McMahon for the $Braf^{CA}$ mouse line, M. Caplan for the *Ksp1.3-Cre* mouse line, the Yale University Animal Genomics Service for pronuclear injections, and Maria Nagy and Art Horwich for technical assistance. *pNI-CD* was a gift from G. Felsenfeld (NIDDK). Many thanks to members of the Xu laboratory for helpful discussions, especially H. Qi, S. Ding, and J. Choi.

## Author Contributions

Conceived and designed the experiments: TX SL JC TN. Performed the experiments: SL JC TN. Analyzed the data: TX SL JC TN MB. Contributed reagents/materials/analysis tools: TX SL JC TN MB. Wrote the paper: TX SL JC TN.

## References

1. Cooley L, Kelley R, Spradling A (1988) Insertional mutagenesis of the Drosophila genome with single P elements. Science 239: 1121–1128.
2. Kleckner N, Roth J, Botstein D (1977) Genetic engineering in vivo using translocatable drug-resistance elements. New methods in bacterial genetics. J Mol Biol 116: 125–159.
3. Ding S, Wu X, Li G, Han M, Zhuang Y, et al. (2005) Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell 122: 473–483.
4. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91: 501–510.
5. Wu S, Ying G, Wu Q, Capecchi MR (2007) Toward simpler and faster genome-wide mutagenesis in mice. Nat Genet 39: 922–930.
6. Bender AM, Collier LS, Rodriguez FJ, Tieu C, Larson JD, et al. Sleeping beauty-mediated somatic mutagenesis implicates CSF1 in the formation of high-grade astrocytomas. Cancer Res 3557–3565.
7. Collier LS, Adams DJ, Hackett CS, Bendzick LE, Akagi K, et al. (2009) Whole-body sleeping beauty mutagenesis can cause penetrant leukemia/lymphoma and rare high-grade glioma without associated embryonic lethality. Cancer Res 69: 8429–8437.
8. Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA (2005) Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. Nature 436: 272–276.
9. Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA (2005) Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature 436: 221–226.
10. Dupuy AJ, Rogers LM, Kim J, Nannapaneni K, Starr TK, et al. (2009) A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. Cancer Res 69: 8150–8156.
11. Keng VW, Villanueva A, Chiang DY, Dupuy AJ, Ryan BJ, et al. (2009) A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. Nat Biotechnol 27: 264–274.
12. Rad R, Rad L, Wang W, Cadinanos J, Vassiliou G, et al. (2011) PiggyBac Transposon Mutagenesis: A Tool for Cancer Gene Discovery in Mice. Science 330(6007): 1104–7.
13. Starr TK, Allaei R, Silverstein KA, Staggs RA, Sarver AL, et al. (2009) A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. Science 323: 1747–1750.
14. Lee T, Luo L (2001) Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. Trends Neurosci 24: 251–254.
15. Pagliarini RA, Xu T (2003) A genetic screen in Drosophila for metastatic behavior. Science 302: 1227–1231.
16. White RM, Sessa A, Burke C, Bowman T, LeBlanc J, et al. (2008) Transparent adult zebrafish as a tool for in vivo transplantation analysis. Cell Stem Cell 2: 183–189.
17. Xu T, Rubin GM (1993) Analysis of genetic mosaics in developing and adult Drosophila tissues. Development 117: 1223–1237.
18. Elick TA, Bauser CA, Fraser MJ (1996) Excision of the piggyBac transposable element in vitro is a precise event that is enhanced by the expression of its encoded transposase. Genetica 98: 33–41.
19. de Wet JR, Wood KV, DeLuca M, Helinski DR, Subramani S (1987) Firefly luciferase gene: structure and expression in mammalian cells. Mol Cell Biol 7: 725–737.
20. Dankort D, Curley DP, Cartlidge RA, Nelson B, Karnezis AN, et al. (2009) Braf(V600E) cooperates with Pten loss to induce metastatic melanoma. Nat Genet 41: 544–552.
21. Dankort D, Filenova E, Collado M, Serrano M, Jones K, et al. (2007) A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. Genes Dev 21: 379–384.
22. Davis IJ, Hsi BL, Arroyo JD, Vargas SO, Yeh YA, et al. (2003) Cloning of an Alpha-TFEB fusion in renal tumors harboring the t(6;11)(p21;q13) chromosome translocation. Proc Natl Acad Sci U S A 100: 6051–6056.
23. Kuiper RP, Schepens M, Thijssen J, van Asseldonk M, van den Berg E, et al. (2003) Upregulation of the transcription factor TFEB in t(6;11)(p21;q13)-positive renal cell carcinomas due to promoter substitution. Hum Mol Genet 12: 1661–1669.
24. Sidhar SK, Clark J, Gill S, Hamoudi R, Crew AJ, et al. (1996) The t(X;1)(p11.2;q21.2) translocation in papillary renal cell carcinoma fuses a novel gene PRCC to the TFE3 transcription factor gene. Hum Mol Genet 5: 1333–1338.
25. Weterman MA, Wilbrink M, Geurts van Kessel A (1996) Fusion of the transcription factor TFE3 gene to a novel gene, PRCC, in t(X;1)(p11;q21)-positive papillary renal cell carcinomas. Proc Natl Acad Sci U S A 93: 15294–15298.
26. Argani P, Antonescu CR, Illei PB, Lui MY, Timmons CF, et al. (2001) Primary renal neoplasms with the ASPL-TFE3 gene fusion of alveolar soft part sarcoma: a distinctive tumor entity previously included among renal cell carcinomas of children and adolescents. Am J Pathol 159: 179–192.
27. Haq R, Fisher DE Biology and clinical relevance of the micropthalmia family of transcription factors in human cancer. J Clin Oncol 29: 3474–3482.
28. Grachtchouk M, Mo R, Yu S, Zhang X, Sasaki H, et al. (2000) Basal cell carcinomas in mice overexpressing Gli2 in skin. Nat Genet 24: 216–217.
29. Sasaki H, Nishizaki Y, Hui C, Nakafuku M, Kondoh H (1999) Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of

Gli2 and Gli3 as primary mediators of Shh signaling. Development 126: 3915–3924.
30. Roessler E, Ermilov AN, Grange DK, Wang A, Grachtchouk M, et al. (2005) A previously unidentified amino-terminal domain regulates transcriptional activity of wild-type and disease-associated human GLI2. Hum Mol Genet 14: 2181–2188.
31. Gorlin RJ (1987) Nevoid basal-cell carcinoma syndrome. Medicine (Baltimore) 66: 98–113.
32. Johnson RL, Rothman AL, Xie J, Goodrich LV, Bare JW, et al. (1996) Human homolog of patched, a candidate gene for the basal cell nevus syndrome. Science 272: 1668–1671.
33. Beauchamp E, Bulut G, Abaan O, Chen K, Merchant A, et al. (2009) GLI1 is a direct transcriptional target of EWS-FLI1 oncoprotein. J Biol Chem 284: 9074–9082.
34. Joo J, Christensen L, Warner K, States L, Kang HG, et al. (2009) GLI1 is a central mediator of EWS/FLI1 signaling in Ewing tumors. PLoS One 4: e7608.
35. Yuan Y, Altman S (1995) Substrate recognition by human RNase P: identification of small, model substrates for the enzyme. EMBO J 14: 159–168.
36. Dassule HR, Lewis P, Bei M, Maas R, McMahon AP (2000) Sonic hedgehog regulates growth and morphogenesis of the tooth. Development 127: 4775–4785.
37. Shao X, Somlo S, Igarashi P (2002) Epithelial-specific Cre/lox recombination in the developing kidney and genitourinary tract. J Am Soc Nephrol 13: 1837–1846.
38. Bosenberg M, Muthusamy V, Curley DP, Wang Z, Hobbs C, et al. (2006) Characterization of melanocyte-specific inducible Cre recombinase transgenic mice. Genesis 44: 262–267.
39. Lechler T, Fuchs E (2005) Asymmetric cell divisions promote stratification and differentiation of mammalian skin. Nature 437: 275–280.
40. Arnold I, Watt FM (2001) c-Myc activation in transgenic mouse epidermis results in mobilization of stem cells and differentiation of their progeny. Curr Biol 11: 558–568.
41. Waikel RL, Kawachi Y, Waikel PA, Wang XJ, Roop DR (2001) Deregulated expression of c-Myc depletes epidermal stem cells. Nat Genet 28: 165–168.
42. Epstein EH (2008) Basal cell carcinomas: attack of the hedgehog. Nat Rev Cancer 8: 743–754.
43. Wang GY, Wang J, Mancianti EH, Jr. Basal cell carcinomas arise from hair follicle stem cells in Ptch1(+/-) mice. Cancer Cell 19: 114–124.
44. Youssef KK, Van Keymeulen A, Lapouge G, Beck B, Michaux C, et al. Identification of the cell lineage at the origin of basal cell carcinoma. Nat Cell Biol 12: 299–305.
45. Littlewood TD, Hancock DC, Danielian PS, Parker MG, Evan GI (1995) A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. Nucleic Acids Res 23: 1686–1690.
46. Wu X, Wu J, Huang J, Powell WC, Zhang J, et al. (2001) Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. Mech Dev 101: 61–69.
47. Uren AG, Mikkers H, Kool J, van der Weyden L, Lund AH, et al. (2009) A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nat Protoc 4: 789–798.